DTIC FILE COPY

(1)

AD-A189 638

AFIT/GOR/ENS/87D-18

A METHODOLOGY FOR FORECASTING VOLUNTARY RETENTION RATES FOR AIR FORCE PILOTS

THESIS

James R. Simpson
Captain, USAF

AFIT/GOR/ENS/87D-18

DTIC
ELECTE
MAR 0 2 1988
S D
H

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

88 3 01 041

AFIT/GOR/ENS/87D-18

A METHODOLOGY FOR FORECASTING VOLUNTARY RETENTION RATES FOR AIR FORCE PILOTS

THESIS

Presented to the Faculty of the School of Engineering

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Operations Research

James R. Simpson, B.S.

Captain, USAF

December 1987

Approved for public release; distribution unlimited

Acknowledgements

I would like to express deep appreciation to several people who contributed to my understanding of the problem and supported me throughout this effort. I would like to thank Major Jon Vetterlein and Captain Don Brazelton of the Officer and Economic Analysis Branch at the Pentagon for helping me understand their model and for providing me with the vital data necessary to build this model. I also wish to thank Major Brian Sutter and Captain Steve Graf of the Personnel Analysis Division at HQ AFMPC for providing the initial motivation for the study, and for sponsoring my work. Thanks is also owed to Major Dana Longino and Major Greg Seidenberger for sharing their knowledge of pilot retention and the airline industry. Most importantly, I would like to thank my advisors, Lt Col Bruce Christensen and Major Joe Litko, for their time, guidance, support, and encouragement. Thank you very much.

James R. Simpson

DTIC COPY INSPECTED 1

Accession For
NTIS GRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification

By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

Table of Contents

		Page
Acknowledgements		ii
List of Figures		v
List of Tables		vi
Abstract		vii
I.	Introduction	1
	General Issue	1
	Specific Problem	3
	Subsidiary Objectives	3
	Scope	4
	Literature Review	5
II.	Data Description and Model Overview	12
	Introduction	12
	Data Description	12
	Overview of the Model	19
	Model Assumptions	20
	Summary	21
III.	Methodology	22
	Introduction	22
	Data Collection	22
	Building the Model	23
	Measures of Performance	28
	Residual Analysis	29
	Corrections for Departures from the Model	31
	Validating the Model	34
	Model Updates	35
	Summary	36
IV.	Findings and Analysis	37
	Introduction	37
	Results of the Initial Model	37
	Results of Variance Stabilizing Techniques	38
	Results of Models with New Predictors	40
	Choosing the Best Model	42
	Validating the Model	44
	Model Updates	47
	Forecasts of Pilot Retention Rates for 1988	49
	Summary	50

	Page

V. Conclusions and Implications 51

Introduction 51
Practical Implications of the Results 51
Recommendations for Refinement 53

Appendix A: Plot of Residuals versus Predicted for the ACOL Model 57

Appendix B: Plot of Residuals versus Time for the ACOL Model 58

Appendix C: Plot of Residuals versus Predicted for the Weighted Least Squares Model 59

Appendix D: Plot of Residuals versus Predicted for the ACOL Model with Transformed Retention Rates 60

Appendix E: Plot of Residuals versus Time and Analysis of Variance Table- Pay Model 61

Appendix F: Analysis of Variance Table and Residual Plots- Job Model 65

Appendix G: Analysis of Variance Table and Residual Plots- Profit Model 68

Appendix H: Analysis of Variance Table and Residual Plots- 1986 Pay Model 71

Appendix I: Analysis of Variance Table and Residual Plots- 1987 Pay Model 74

Appendix J: Analysis of Variance Table and Residual Plots- 1987 Profit Model 77

Appendix K: Data Used in the Pay and Profit Models 80

Appendix L: Plots of Actual and Predicted with Prediction Intervals- Pay Model 81

Appendix M: Plots of Actual and Predicted with Prediction Intervals- Profit Model 86

Bibliography 91

Vita 93

List of Figures

Figure Page

1. Prototype Regression Function Using Indicator Variables 26

2. Residual Plot of Decreasing Error Term Variance 30

List of Tables

Table		Page
I.	Pilot Retention Prediction Intervals Using Forecasted Independent Variables	36
II.	Results of Weighted Least Squares	39
III.	Model Fit Results	43
IV.	Pay Model Forecasts for 1986	45
V.	Results of the Pay Model with Different Forms of the Airline Hires Predictor	46
VI.	Pay Model Forecasts for 1987	47
VII.	Pay Model Update Results	48
VIII.	Profit Model Update Results	49
IX.	Pay Model Forecasts for 1988	50

Abstract

The purpose of this study is to develop a model that more accurately forecasts voluntary retention rates in the short term for Air Force pilots. Specifically, the model consists of appropriate and available predictors used to compute one year ahead forecasts of voluntary retention rates for Air Force pilots with seven through eleven years of service. Previous and existing military retention models were reviewed to study appropriate predictors and methodologies.

The types of predictors collected for study were indicators of the strength of the economy, indicators of the growth of the airline industry, and indicators of the relative wage difference between the military and the civilian labor force. Classical regression analysis was used to predict the pilot retention rates on the basis of the predictor variables studied. Because the dependent variable is a ratio, bounded above and below, transformations and weighted least squares were implemented in an effort to stabilize the error term variance. The most successful variance stabilizing technique was a logarithmic transform of the pilot retention rates.

The criteria established for selecting the best model were model performance, prediction potential, and explanatory

significance. The best model included the following independent variables: indicator variables for the year of service groups, a variable for the annual number of new airline pilot hires, the unemployment rate, and a pay compensation measure. The unemployment rate and the pay compensation measure were significant leading indicators of pilot retention rates, and therefore were lagged variables. Thus, estimates were required only for the airline hires predictor in order to forecast pilot retention rates. An alternative model was proposed which included the indicator variables, airline hires, the unemployment rate, and corporate profits.

Validation tests were performed on the best model for years 1986 and 1987. In each test, the 90 percent prediction intervals covered the actual pilot retention rate for each year of service group. Among the recommendations provided to improve the accuracy of the pilot retention rate forecasts was to improve the accuracy of the airline hire forecasts and to find other significant, leading indicators of pilot retention.

A METHODOLOGY FOR FORECASTING VOLUNTARY RETENTION RATES OF AIR FORCE PILOTS

I. Introduction

General Issue

Retention of Air Force personnel has always been an important and challenging objective. In particular, pilot retention is crucial because of the additional training costs and, more importantly, the time needed to train an experienced pilot. If the Air Force intends to meet future force capability requirements, it must be able to replace in a timely manner the pilots who leave the service. Proper replacement can only be achieved by anticipating the number of pilots that will leave. Thus, there is a need to accurately estimate future pilot retention.

According to Major Brian Sutter, Chief Rated Analyst for the Air Force Personnel Analysis Division, the Air Force senior leadership has requested more accurate forecasts of retention rates. Specifically, they need a model that can be used to compute forecasts of voluntary retention rates for certain year groups (12). Voluntary retention rates are the percentage of pilots without an Active Duty Service Commitment who voluntarily remain in the service.

Air Force leadership is especially interested in accurate forecasts of voluntary retention rates for pilots with seven through eleven years of service. Under Air Force regulations, pilots have an Active Duty Service Commitment until their seventh year. Historically, the voluntary retention rates for pilots with more than eleven years of service have been very high and consistent. Therefore, the primary focus in voluntary retention behavior is on pilots with seven through eleven years of service.

The Air Force Retention Planning Committee met in 1986 and discussed measures of retention used by the Air Force. Lieutenent Colonel Katnik, Branch Chief of the Officer and Economic Analysis Branch at the Pentagon, submitted a paper to the committee suggesting that four measures be used to tie retention to force capability. These measures include expected retention, required retention, objective force retention, and actual retention (8:2). The expected retention is the level of retention forecast by the Air Force models. In order to accurately tie these measures to force capability, accurate forecasts of retention must be used.

The model used by the Air Force to forecast pilot retention is the Officer Personnel Analysis System. This system consists of three component models used and maintained by the Directorate of Personnel Plans, Pentagon, to determine Air Force personnel force requirements. The three components include the Compensation model, the Econometric Adjustment model and the Inventory Projection model. The Compensation

model computes the Annualized Cost of Leaving the Service (ACOLS) by aeronautical rating and by year of service. ACOLS is a fairly complex variable which measures the relative difference between lifetime earnings of military officers and the earnings of their civilian counterpart. The ACOLS measure is used as an input to the Econometric Adjustment model. This model computes the expected changes in the retention rates using historical relationships between ACOLS, unemployment rates, hiring by the major airlines, and the retention behavior of Air Force officers. These expected changes are then used as inputs to the Inventory Projection model to forecast retention rates and project force requirements. The Econometric Adjustment model has not been updated since 1983.

Specific Problem and Research Objective

The Econometric Adjustment model of the Officer Personnel Analysis System requires updating and improvement. The purpose of this research is to develop a model that more accurately forecasts voluntary retention rates by year group in the short term for Air Force pilots. Specifically, the model will use appropriate and available predictors to compute forecasts of retention rates for Air Force pilots with years of service seven through eleven.

Subsidiary Objectives

The sub-objectives that must be attained to completely attain the research objective are the following:

1. Determine the specific purpose or use of the model.

2. Determine who will be using the model.

3. Determine what models currently exist.

   a. Determine if similar civilian or foreign models exist.

   b. If similar models do exist, determine how they can be modified to address this specific problem.

   c. If similar models do not exist, determine the problems people have encountered trying to develop them.

4. Determine what type of model should be used.

5. Determine what use historical rates will have in predicting retention rates.

   a. Determine the reliability of the data.

   b. Determine how the data is defined.

6. Determine which economic factors influence retention.

   a. Determine which factors are used in similar models.

   b. Determine what data are available.

7. Determine how the model will be verified and verify the model.

8. Determine how the model will be validated and validate the model.

Scope

The problem of forecasting retention rates for the Air Force is too large to address in a single thesis research effort. Thus, the scope will be narrowed to include the following:

1. Air Force pilots--the eligible population will include all line officers in the grade of lieutenant colonel or below, not suspended from flying duties.

2. Short term forecasts (e.g. one year ahead).

3. Year groups seven through eleven.

4. Voluntary retention rates.

With the specific objective of forecasting pilot voluntary retention rates in the short term, a search of work done in in this area was conducted to determine appropriate predictor types and methodologies.

Literature Review

This section is a review of some of the work documented in the field of military retention modeling. The focus of the review is on the retention models currently used by the Air Force and the retention models developed for, but not currently used by the Air Force. The modeling techniques and the factors used as inputs to the models are discussed.

The analysis system used by the Analysis Division of the Directorate of Personnel Plans is the primary personnel analysis tool in the Air Force. The three model system ages the Air Force by projecting retention (both voluntary and involuntary), accession, promotions, flying suspensions, and the flight training turnover. Voluntary retention of Air Force officers is projected using the first two models, the Compensation model and the Econometric Adjustment model (15).

Forecasts of voluntary retention are obtained by adding estimated future changes in the retention rates (called delta retention rates) to the previous year's rates. These delta

retention rates are computed using a general linear model. A logistic transformation is made of the delta retention rates so that the assumption of constant variance throughout the predictions is maintained. Ordinary least squares regression is used to estimate the parameters of the model which include an intercept term and coefficients for each of the predictors. The predictors are the number of major airline hires, the unemployment rate, and the ACOLS measure (15).

Retention rates are computed for many different groups of officers. Officers are broken into classes by component (regular or reserve), source of commission, grade, aeronautical rating, and years of service. The annual retention rates are then converted to monthly rates using historical seasonality (15).

In addition to the Officer Personnel Analysis System, several models do exist that address retention of Department of Defense personnel. These models offer insight into the reasons why people decide to stay in or leave Federal service. Some models are concerned with the effect certain retirement and personnel policies have on peoples' attitudes toward service and on their decision to stay or leave. Some models show how the economic factors affect retention of personnel. Other models primarily focus on the wage differences between Federal and non-Federal employees as a determinant of retention. Each of these models will be

discussed. Included in each discussion will be the purpose, the methodology, and the various inputs and outputs to the model.

A model developed by Gotz and McCall of RAND Corporation calculates the probability that an Air Force officer will voluntarily remain in the service based on a given set of retirement, compensation, and promotion policies. According to Gotz (7:1), this model is a stochastic dynamic program with the purpose of assessing the retention implications of alternative compensation and personnel policies. These policies are inputs to the model. Voluntary retention rates are output by fiscal year, rating, source of commission, years of service, component, and grade (7:2).

The voluntary retention rates are determined in the dynamic program by finding the individual officer's optimum time to leave the military. According to Gotz, this optimum time occurs when the individual's expected present value of pecuniary and non-pecuniary returns are maximized (7:1). The parameters of the model are estimated by maximum likelihood. A distribution of the taste for military service is included in this model. Tastes are assumed to follow the extreme value distribution for maxima. This distribution is skewed to the right, meaning it has a long right-hand tail. Gotz chose this skewed distribution for the following reason. While we may expect to observe officers who place almost infinite value on remaining in the service, it is unlikely

that those who value not being in the Air Force in the same amount would have joined in the first place (7:18).

Saving and DeVany developed a general model of the Air Force manpower market. Their approach was to develop a stochastic process model of both the accession and retention markets of Air Force enlisted personnel (1:1). They treated the problem as a queueing process by viewing the allowable force as the number of servers in the process and the mean length of stay as the service time. The retention portion of the model will be the focus of this discussion. Saving and DeVany developed a utility maximizing model which yields the optimal distribution of total working life between military and civilian alternatives (1:3). The mean length of stay depends on the relative wages (military versus civilian), minimum quality standards of new enlistees, and the minimum enlistment period (2:10).

Since working with Devany on the manpower model, Saving has developed a more extensive retention model that considers both the occupational and individual characteristics as well as policy and force management factors. The primary purpose of the model is to determine the retention of enlisted personnel within Air Force Specialty groups. This model determines the probability that an airman will reenlist, given the airman's vector of attributes (3:5). Because the decision to reenlist is a binary one, the shape of the response function (for retention rates) will frequently be curvilinear (10:361). This function is often shaped like a

tilted S, and has asymptotes at zero and one. Transforming this function by means of a cumulative normal distribution into a linear function is called a probit transformation (10:366). The transformed probit model can easily be extended into a multiple regression model for use in forecasting. Saving used the probit model with the airman's attributes as inputs. The attributes he used included the following:

academic education level;

race;

Armed Forces Qualification test scores;

number of dependents;

sex;

marital status;

real military compensation (the present value of a 4-year income stream);

the employment rate;

reenlistment bonus;

civilian wage (3:7,8).

The parameters are estimated using maximum likelihood. Saving discusses his hypothesis of the influence of each of these input variables on the retention rate and finishes by analyzing the empirical results.

Cromer and Julicher developed a model to describe Air Force pilot retention rates. Their objectives were to build a model based on economic conditions, determine the model's predictive potential, and determine the significance of airline hires on pilot retention (5:4). In an attempt to

find the "best" model, they used three methods: factor analysis, stepwise multiple regression, and multiple regression with lagged retention rates.

In each model they started with the same set of sixteen different economic factors. All but four of these factors were obtained from the Business Conditions Digest, a monthly report by the Bureau of Economic Analysis. Each of the factors from the digest are classified as leading, coincident, or lagging according to the timing of the peaks, troughs, and turns in the time series relative to the business cycle (5:16). The factors were originally selected because they were believed most likely to have an effect on individual behavior or because they are indices for interpreting current, or predicting near-future business conditions (5:16). Some of the factors include the Consumer Price Index (CPI), white collar unemployment, the average prime rate, and the lag of real military pay with respect to CPI (5:17-21).

The results of the factor analysis method showed that no model accurately described retention rates (5:28). Cromer and Julicher then used unlagged retention rates and stepwise multiple regression. Stepwise multiple regression is a method that considers each economic factor in the presence of all other factors and adds a factor to the model if it significantly contributes to the model by describing retention. The results of the stepwise regression showed that six economic factors were significant as unlagged series

and the model described all of the variability of the retention rate data (5:31).

Using lags of six months and twelve months for the retention rates Cromer and Julicher used stepwise regression to build two final models. The six month lag model contained two significant economic factors and performed well in describing retention rate data. The twelve month lag model contained four factors and also did a good job of describing the retention data (5:33). Cromer and Julicher conclude that the unlagged stepwise regression model is the "best" model in terms of describing the retention rate data. They do not suggest the use of the model to predict retention because not enough data were available to build and validate a forecasting model (5:51). They also state that airline hires did not appear as a significant factor in any of the models (5:51).

The methodologies and inputs discussed in the above review provided the guidance for the methodology selection and choice of predictors to be studied in this research effort. The Econometric Adjustment model, used by the Air Force to predict pilot retention rates, was used as the basis for the model development. The types of predictors used in this effort are representative of the inputs used in the models discussed above. The relation between these types of predictors and pilot voluntary retention are discussed in the next chapter.

II. *Data Description and Model Overview*

Introduction

The input data collected for study, in addition to the methodology selected, are important factors in developing an accurate forecasting model. This chapter begins with a discussion of the motivation for the types of data collected, followed by a description of each of the series collected. An overview of the methodology and the accompanying model assumptions are then presented.

Data Description

Based on the previous work done in retention modeling and the types of inputs used in those models, an effort was made to obtain similar, appropriate data for this research effort. This section contains a description of the types of explanatory inputs relating to the military voluntary retention behavior. Relationships between pilot retention, the strength of the economy, the growth of the airline companies, and the relative wages of pilots to their civilian counterparts will be discussed. The organizations that provided data for this study will be acknowledged. The section will conclude with a description of each of the variables used in developing the model.

Pilot Retention and the Economy. Pilots in the position to make a decision about their future, those who may consciously choose between staying in or leaving the military, are most likely concerned about the strength of the

economy. Just wanting to leave the military usually is not enough reason to cause pilots to resign. Pilots, as rational decision makers, are concerned about the availability and appeal of jobs in the civilian labor force. Plentiful, quality jobs are related to the strength of the economy. When the economy is healthy, civilian jobs are more attractive to Air Force pilots. In a strong economy versus a weak economy, civilian jobs are more secure and financially rewarding. Major Gentile, formerly of the Officer Branch of the USAF Retention Division, studied the relationship between pilot retention and the economy. He reported that the pilot retention rates have been strongly correlated with the strength of the economy (6:viii). Using the white collar unemployment as a measure of the strength of the economy, he observed that the trends of the two series from 1978 through 1982 are almost mirror images (6:24).

Recently, several models using economic factors as explanatory variables have been designed by analysts to help study the behavior of military retention rates. Saving developed a model to study the behavior of enlisted personnel. He believes that retention decisions of Air Force enlisted personnel have always been significantly affected by economic factors (3:1). Cromer and Julicher developed a model of pilot retention behavior based on economic indicators. They applied the utility theory to pilot career decisions by asserting that pilots' stay/leave decisions are dominated by their own economic perceptions, with the actual

economic environment exerting influence (5:15). These analysts have found economic factors to be statistically related to military retention rates over the past ten years.

Pilot Retention and Relative Wages. Whether pilots are considering leaving the military to fly for the airlines or to work in a non-flying job, they are interested in the wages of civilian jobs relative to their military wages. The Annualized Cost of Leaving (ACOL) measure, developed in the Compensation model, uses the difference between career earnings in the military and career earnings as a civilian in a similar occupation to determine the optimal time to leave the service. The ACOL measure is used as a predictor in the Econometric model to forecast the expected changes in retention rates (Vet). Saving's enlisted retention model uses military pay compensation as a predictor. His model uses inputs from military and civilian streams of earnings (Sav 85, p.7-9). Both models show a positive correlation between the relative wage difference and miltary retention. If the military pay increases are not keeping pace with the raises in the civilian labor force, military retention declines.

Pilot Retention and the Airlines. Pilot retention is affected by the lure of the airline industry. According to Major Longino, of the Officer Branch at the USAF Retention Division, seventy-five percent of all Air Force pilots intending to leave the service plan to fly for the airlines (9). For pilots to actually separate for this reason, the airlines must be hiring. Major Gentile reports that there is

a direct correlation between airline hiring and USAF pilot retention, so that when the airlines hire, USAF pilot retention suffers (6:viii). The combination of these outside pressures is adversely affecting the pilot retention rates, which have been dropping since 1984.

Lieutenant Colonel Rhodes, in his historical analysis of USAF pilot retention, reports that a booming economy combined with plentiful airline jobs on the outside is the primary external reason for pilot losses (11:8). Indicators of the strength of the economy, growth of the airline industry, and the relative wage difference between military and civilian workers, were sought for study as inputs to the retention prediction model.

Data Sources. The data used in this study were provided by several sources. The voluntary retention rates for Air Force pilots by year of service from 1977 through 1987 were provided by the Officer Branch of the USAF Retention Division, at the Headquarters Air Force Manpower Personnel Center. The ACOL measures, the airline hires, and the pay compensation data were provided by the Personnel Analysis Division of the Directorate of Personnel Plans, the Pentagon. Other data used in this study were economic indicators obtained from the Business Conditions Digest, a monthly periodical published by the Bureau of Economic Analysis.

Model Variables. The following variables were selected for study in the model. Six predictors, representing the three types of data, were studied. The list contains a

description of each data series, the reasons for selecting the variable, and the expected influence that each variable will have on the pilot retention rates.

Pilot Voluntary Retention Rates - The data provided by the USAF Retention Division are annual pilot voluntary retention rates by year of service. These rates were first available in 1977, so eleven years of data are used in this study (through 1987). The data is recorded each fiscal year. To avoid double counting, a pilot's year of service is defined as the number of years of service completed at the beginning of a particular fiscal year.

Annualized Cost of Leaving (ACOL) - The ACOL measure is also calculated by year of service for each fiscal year. The actual data obtained from the Analysis Division at the Pentagon are changes in the ACOL for a pilot with a certain number of years of service in a certain fiscal year. The reason for considering this measure for use as an explanatory variable is that it indirectly measures the individual's taste for military service. The ACOL measure and voluntary pilot retention rates are expected to be positively correlated. As the change in ACOL increases, voluntary retention is expected to increase.

Pay Compensation - The pay compensation ratio is also a measure of the relative difference between military and civilian earnings. This measure is a ratio of a military pay index to a civilian pay index. The base year for this ratio is 1972. In that year the relative earnings for similar jobs

between the military and the civilian labor force is assumed to be equal, so the ratio equaled one. Each year the military pay index changes by the percentage increase in military pay. The index for civilian pay is measured using the Employment Cost Index (ECI). The ECI is a quarterly measure of the average change in the cost of employing labor. This index includes wages, salaries, and employer costs for employee benefits and covers over 400 occupations in the private nonfarm and public sectors (about 70 percent of military jobs). The ECI is not affected over time by changes in the composition of the labor force (14:103). As an explanatory variable, pay compensation should be positively correlated to voluntary pilot retention rates. If the ratio increases, pilot retention rates should also increase.

Airline Hires - The data for the number of airline hires are compiled by the Future Aviation Professionals of America (FAPA). The hires include all new hires by companies flying jet aircraft. This group includes major, national, and turbojet companies. The number of new hires for regional airline companies, which fly propeller-driven aircraft, are not included.

There is a possibility that the number of new hires for jet aircraft companies has been slightly inflated since 1985. Due to the shortage of pilots in the industry, airlines have begun recruiting pilots from other airlines to fill vacancies. It is possible that double counting is taking place because pilots are transferring between airline

companies. More accurate counts of new hires are not currently available. The expected correlation between hires and retention is negative. As hires increase, pilot voluntary retention should decrease.

Unemployment Rate - This series is based on data collected in household surveys conducted each month by interviewers of the U.S. Department of Commerce. The unemployment rate is the ratio of the number of persons unemployed to the civilian labor force (13:9) This economic indicator is inversely related to broad movements in aggregate economic activity (13:10). If the unemployment rate is increasing, the retention rate for pilots should also increase.

Corporate Profits - Corporate profits is the income of corporations organized for profit plus the income of mutual financial institutions that accrues to residents, measured before profits taxes. Profits tax includes Federal, State, and local taxes on corporate income (13:26). The current-year profits are then converted to constant (1982) dollars. This measure is considered by Business Conditions Digest to be a leading economic indicator. A negative correlation between pilot retention and corporate profits is expected. If corporate profits are decreasing, pilot retention rates should increase.

Help-Wanted Advertising in Newspapers - This series is an index (1967=100) that measures employers' demands for labor. The index reflects the relative level and monthly

change in the number of job openings resulting from vacancies in existing jobs or the creation of new jobs. The data are based on the daily volume of help-wanted ads published in the classified section of one newspaper in each of 51 sample cities. Each city represents a major labor market area as defined by the Bureau of Labor Statistics (13:9). The reason for considering this index is to reflect the availability of jobs in the private sector for pilots. As advertising increases, pilot retention rates are expected to decrease.

Overview of the Model

The choice of methodologies was based on the requirement that explanatory variables must be used as inputs to the model. A modeling technique which can use the relation between these explanatory variables and the pilot retention rates is general linear regression. Bowerman states that classical regression analysis is a very useful statistical technique that can be used to predict a dependent variable (retention rates) on the basis of one or more independent (explanatory) variables (4:393). The general linear regression model can be defined as follows:

$$Yj = B0 + B1Xj1 + \ldots + BKXjK + ej \qquad (1)$$

where

Yj is the value of the dependent variable in the jth trial
B0,B1,...,BK are parameters to be estimated
Xj1,...,XjK are known constants, the value of the independent variables in the jth trial
ej are error terms

The parameters of the model are estimated using the

method of least squares, a technique designed to minimize the sum of the squares of the error terms. The residuals are equal to the difference between the actual response (pilot retention rate) and the response estimated by the regression equation. The technique is called general *linear* regression because the function is linear in the parameters, meaning that no parameter appears as an exponent, or is multiplied or divided by another parameter (10:31).

Model Assumptions

The assumptions of the pilot voluntary retention model include the assumptions of the general linear model. The error terms are assumed to be random variables with mean of zero and a constant variance. The errors are also assumed to be uncorrelated with each other.

Because the objective of this effort is to make retention rate predictions, interval estimates must be made and statistical tests must be performed. Therefore, an assumption is made about the functional form of the distribution of the error terms. The standard assumption is that the error terms are normally distributed. Under this assumption the error terms are not only uncorrelated, but necessarily independent. The normal error assumption also implies that the dependent variable, the pilot retention rate, is also normally distributed (10:49). Hypothesis tests will be performed to verify the general linear model assumptions.

An assumption is made that the voluntary retention data and the data used for the explanatory variables are accurate. The retention rates used to develop the model are actual population rates, not sample rates. Thus, if the numbers are accurate, they represent the true voluntary retention situation for pilots for fiscal years 1977 through 1987. The same assumption applies to the data for the independent variables.

Summary

The types of predictors considered for study in this model were economic indicators, airline industry growth indicators, and relative wage difference indicators. Specific data series were chosen and collected to represent these three categories. Each series was studied for its logical contribution to the model. Regression analysis was chosen as the methodology for incorporating these predictors into the model with the intention of forecasting pilot retention rates. The model development is discussed in the following chapter.

III. Methodology

Introduction

The procedure for developing a model to more accurately predict pilot retention rates includes the following steps: 1) collecting appropriate and accurate data, 2) building the initial general linear regression model, 3) performing diagnostic tests, and 4) refining and testing the model until all assumptions are met and the best model is identified. Each step is discussed in this section. In the process of model development, some observations were made concerning the difficulties in generating a pilot retention rate model. Some of the lessons learned are discussed to provide insight to those who plan to work with pilot retention rate models.

Data Collection

Data collection is the first step in the model development process. The data should, as best as possible, represent the real world situation. For instance, the pay compensation data should ideally represent the actual differences in earnings between pilots in the Air Force and civilians in similar occupations. The actual pay compensation data collected is average change in earnings between the military and the civilian labor force (using the Employment Cost Index). So, the actual ratio used is not ideal, but representative of the relative earnings. For the

purposes of this study, the data used in this model is assumed to be accurace.

The predictors should be explanatory and should logically relate to the pilot retention rates. The purpose of building the model is to determine whether the relation also exists statistically. To be statistically related, the variability of the predictors should coincide with the variability of the pilot retention rates. Because the ultimate objective of this research effort is to predict the pilot retention rates one year into the future, the accuracy of the prediction will be the greatest if the predictors are all leading indicators of pilot retention. For example, if the objective is to predict the pilot retention rates for fiscal year 1987, it is preferred to have all the predictors as known constants for years 1986 or earlier. Otherwise, forecasts of the predictors must be used to forecast pilot retention rates.

Building the Model

The second step in model development is building an initial general linear model and applying regression analysis to the model. The SAS statistical analysis package was used extensively to assist in building the initial model, performing statistical tests for diagnostic purposes, and to help find the best model. This section will include a discussion of the actions taken to build the best general linear model. Criteria were established based on model

performance, prediction potential, and explanatory significance to select the best model. These criteria are discussed in the following chapter.

The initial model was designed after the Econometric Adjustment model used by the Analysis Division at the Pentagon. The initial independent variables are the same in type as those actually used by the Pentagon. The independent variables are the ACOL measures (with different variables for each year of service group), the numbers of airline hires, and the unemployment rates. The first model contained each of these variables as an unlagged series.

In addition to these predictor variables, the model also contained indicator variables for the year of service groups. This technique is used to account for the pilots' taste for military service. Warner describes this pattern in retention rates.

> ... there should be a natural tendency for retention rates to rise with term of service (t). This tendency is separate and distinct from any increase in the financial incentive to stay and is due to the fact that in early terms of service the retention decision-making process serves to sort out those who like military service from those who don't. As this sorting process proceeds, the cohorts of personnel who stay will... [consist of] people who, on average, have a higher taste for military service and hence higher retention rates (16:3).

With the addition of the indicator variables to the regression function, the retention rate equations now are different for each year of service group (for years seven

through eleven). These indicator variables actually change the intercept of the equation, so separate equations are generated for each year of service group. The lines are displaced by an amount equal to the indicator parameter value. For example, with the indicator variables for each year of service and a single predictor variable, the regression function would include the following:

$$Yj = B0 + B17YOSj + B28YOSj + B39YOSj + B410YOSj + B5Xj + ej \qquad (2)$$

where

Yj	is the pilot retention rate for a specific year of service in year j
B0,B1,...,B5	are the parameters to be estimated
XYOSj	are the indicator variables in year j
Xj	is the value of the predictor in year j
ej	is the error term

The indicator variable for pilots with eleven years of service is implicitly captured in the equation. The intercept term for this year of service is estimated by B0. Figure illustrates a prototype of a regression function using indicator variables for each year of service. The unemployment rates and the airline hires in a given fiscal year are consistent for each year of service group. This situation results in identical slopes for each of the five year of service group equations, similar in concept to figure 1.

The ACOL measure is the only predictor which is computed for each year of service group. If the ACOL measures and the

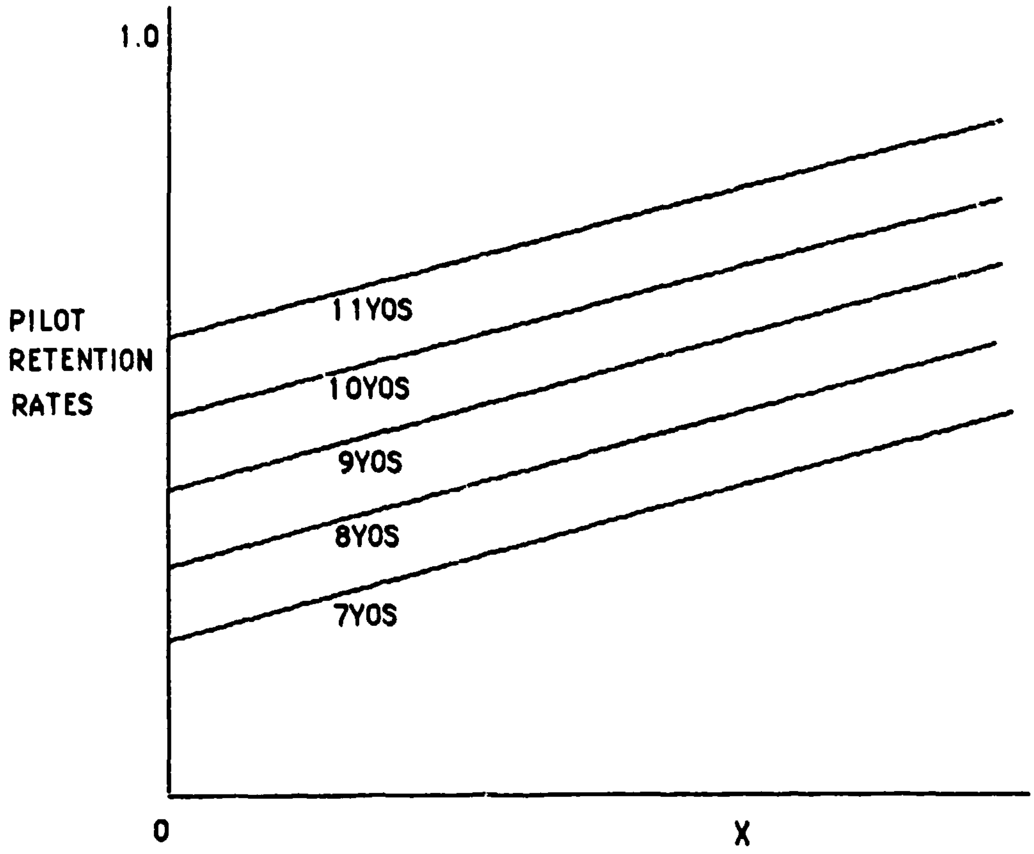


Figure 1

indicator variables for each year of service group are determined to significantly contribute to the model, then both the intercept and the slope of the function may change for each year of service group.

Each of the independent variables used in the model are standardized before they are input into the regression equation. The data is standardized by dividing each value of a particular data series by that series' sample standard deviation. Once each series is standardized, the effect each

independent variable has on the response variable is reflected in the magnitude of the regression coefficients. Also, the data is standardized to keep the magnitude of the coefficients in a reasonable range.

The following procedure was used to determine if independent variables should be lagged or unlagged. Each variable used in the model was first input as an unlagged series. Hypothesis tests were performed to determine the contribution of each independent variable to the model. Then, each independent variable was lagged separately, keeping the other variables unlagged. Because data were available for each of the predictors prior to 1977 (the first year of the pilot retention rate data), no loss of observations occurred. The hypothesis tests were performed on the predictors once again to determine their significance. The results of the models were compared in terms of the t-statistic for the hypothesis test for significance of the parameter estimates. Decisions were made to keep or drop variables based on a critical level of significance of 0.05.

If a particular variable was significant both when lagged and when unlagged, the series which produced the best overall predictive model was used. The significant lagged series carried more weight as a predictor than an unlagged series of equal significance, because the lagged series was statistically a leading indicator of pilot retention. The leading indicators were lagged one year, so estimates of

these indicators did not have to be made when forecasting pilot retention rates one year ahead. If the series was not equally significant as a lagged and unlagged series, then the width of the prediction intervals generated by the two separate models were compared and, if all the model assumptions were maintained, the model generating the smallest prediction interval widths was selected.

Measures of Performance

The statistic used to measure the ability of a set of independent variables in a model to proportionately reduce the total variation in the response variable is the coefficient of multiple determination, denoted by R2. The R2 ranges from zero to one, with a value of one indicating a perfect fit. Adding more independent variables to the model can only increase R2. It is widely accepted that a modified measure, called the adjusted R2, be used to compare models with different numbers of independent variables. The adjusted R2 may actually become smaller when another independent variable is introduced into the model. The mean square error (MSE) is also a measure of the ability of a set of independent variables to reduce the variation of the response variable. The MSE is defined by the following equation:

$$MSE = SSE / (n - p) \qquad (3)$$

where

SSE = error sum of squares
n = number of observations
p = number of parameters estimated (number of predictors + 1)

The MSE can also increase as more predictors are added to the model. The R2, adjusted R2, and MSE were considered, in addition to the widths of the prediction intervals, as measures of model performance.

Residual Analysis

Diagnostic tests were performed to check the validity of the general linear model assumptions by evaluating the residuals. The residuals were studied to examine three possible departures from the general linear model. The possible departures were lack of constant variance, lack of normality, and lack of independence of the error terms. Statistical tests, in addition to graphic analysis of the residual plots, were performed to check for these departures.

The possibility of nonconstant error term variance was first addressed. A plot of the residuals against the predicted values of the retention rates is helpful to study whether the variance of the error terms is constant. Figure 2 is an example of the residual plot when the error term variance decreases with increasing values of predicted variables. This type of deviation most likely would be a problem with this model because the dependent variable is a rate, and most of the data is between 0.6 and 1.0. The

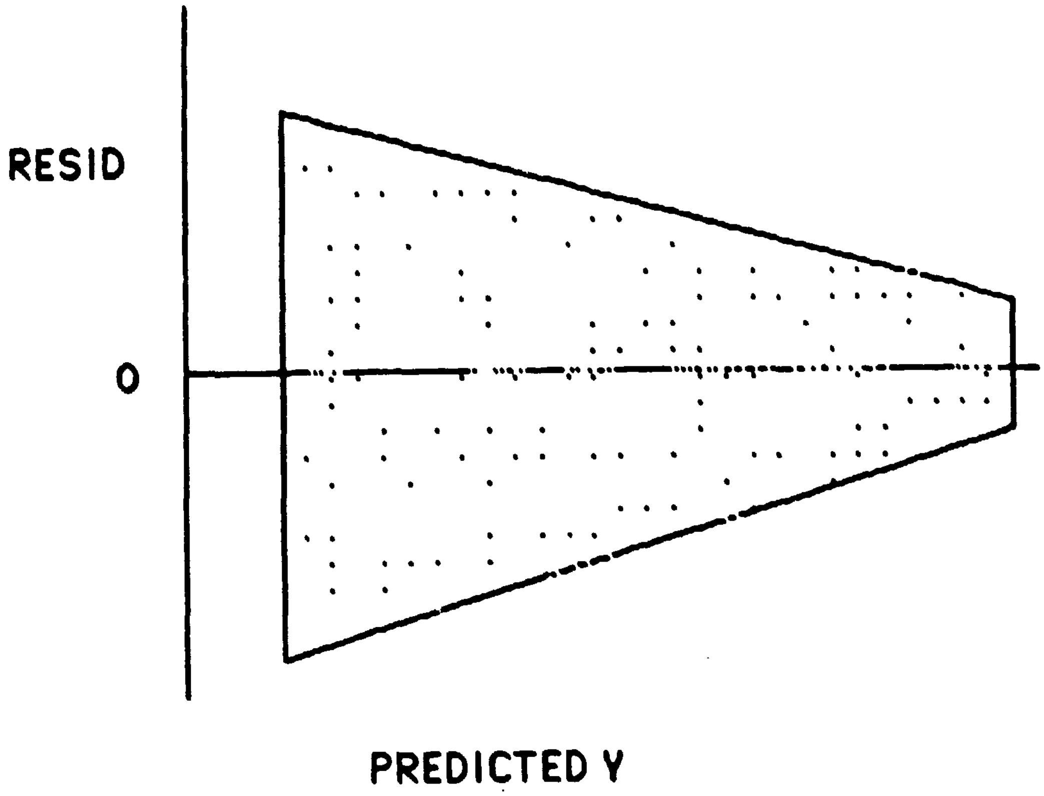


Figure 2

variance of the observations close to 1.0 is expected to be smaller than the variance of observations close to 0.6.

The second possible departure from the model is lack of normality of the error terms. It should first be noted that small departures from normality should not cause problems with the model. The normality of the error terms can be studied graphically by preparing a normal probability plot. The residuals are plotted against their expected values when the distribution is normal. A plot which is almost linear suggests agreement with normality (10:118).

The possibility of lack of independence of the error terms was studied. The purpose of the study was to determine if the error terms are correlated over time. The residuals were plotted against time to study whether a pattern existed. As an additional check, the Durbin-Watson statistic was calculated for each of the models. This test is designed for lack of randomness in the residuals. The value of the statistic is close to two if the error terms are uncorrelated.

Corrections for Departures from the Model

If the general linear model assumptions are not satisfied using a particular model, remedial measures must be taken to correct for the problem or the model must be abandoned. The appropriate measures for each of the three problems are discussed, with emphasis on correcting for nonconstant variance, because it is the main concern of this model. The approach to correct for correlation of the error terms is to add one or more independent variables to the model or to use transformed variables. Transformed variables are also suggested when large deviations from normal error terms exist.

Several techniques are available that help stabilize the error term variance. Weighted least squares is a method used to obtain parameter estimates that often corrects for nonconstant error term variance. Transformation of the

dependent variable can also be effective in stabilizing the error term variance.

Weighted least squares is a variance stabilizing technique that assigns weights to each observation. The weight for an observation is the inverse of the observation's error term variance. The pilot retention rates are proportions of individuals making a decision to voluntarily leave the service. Each individual in the population is making a decision to stay in or leave the service. If each decision can be considered to be a Bernoulli trial, the distribution of the population can be assumed to be binomial with proportion p. The variance of the point estimator p is equal to the following:

$$\mathrm{Var}\ (p) = (p * q) / n \tag{4}$$

where

p = the pilot retention rate
q = the pilot loss rate (1.0 - p)
n = the population size for each observation

The weight for each observation is then the inverse of the variance:

$$\mathrm{weight}_j = n_j / (p_j * q_j) \tag{5}$$

for each jth observation

These weights make sense intuitively for two reasons. First, the weights place more emphasis on the observations

with larger populations. Second, the retention rates close to one, which have less variance, also have larger weights.

A second technique available to help stabilize the error term variance is the dependent variable transformation. When the dependent variable is a rate, the appropriate transformation is a logarithmic transform of pilot retention rates. In this case, the pilot retention rates are all above 0.6. The upper bound (1.0) is the only bound that is constraining the rates. The appropriate transformation is defined by the following equation:

$$\text{TRANSRET} = -\,\text{LN}\,(\text{UB} - \text{P} + \text{DELTA}) \qquad (6)$$

where

TRANSRET	= the transformed pilot retention rate
P	= the original pilot retention rate
UB	= the upper bound (1.0)
DELTA	= a small constant (0.01)

The constant, delta, is determined by trial and error. Four different values were used and the one providing the most consistent error terms was used. The constant 0.01 did the best job with this particular data set, but the impact of this selection on the variance of the error terms is small. The only constraint, using a delta of 0.01, is that all retention rates must be less than 0.99. If rates greater than 0.99 exist, a smaller delta must be used.

The remedial measures discussed above were instrumental in developing the models in this study. Choosing the best model not only requires measuring the performance of each

model, but also requires careful analysis of the residuals and correcting for departures, if possible, once they are detected. The results of the model development process will be discussed in the following chapter.

Validating the Model

The process of developing and choosing the best model required statistical tests on the model assumptions as well as tests of the statistical relationship between the dependent and independent variables. In this sense, the model was being verified in the development stage. The best model was selected because it is expected to do the best job of predicting pilot retention rates. The validation process is necessary to determine whether the best model accurately predicts pilot retention rates.

The statistical model was developed using data from fiscal years 1977 through 1985. The data for 1986 and 1987 were intentionally withheld for validation purposes. Validation tests were performed only on the best model. In order to predict the retention rates for these two years, the following data were required: 1) the regression coefficients for each independent variable, 2) the actual data for any independent variables lagged one year or more, and 3) the forecasts for unlagged independent variables.

The values of the independent variables were input into the model and prediction point estimates were computed. In addition, 90 percent prediction intervals were provided.

Each prediction interval should cover the true retention rate with a probability of 0.90.

The independent variables whose series are significant to the model as unlagged, or lagged less than one year, required estimation in order to predict the pilot retention rates. In some cases, these forecasts of independent variables were given as point estimates and in other cases as a range of values. If a range was specified for the predictor estimate, prediction intervals for the pilot retention rates were computed by running the particular model two separate times, using the high and low value of the forecast for the predictor. The retention rate interval was then built by using the smallest values of the two lower bounds and the highest value of the two upper bounds (see Table I). Obviously, the interval width will increase with range estimates versus point estimates, but the probability that the interval will cover the actual retention rate will also increase.

Model Update

Following successful completion of validation testing, the model was updated using data from 1986 and 1987. By adding these two years of data to the model, the input database increased significantly (approximately twenty percent). The regression coefficients were compared against the original model to determine if any coefficients significantly changed. The general linear model assumptions

TABLE I. PILOT RETENTION PREDICTION INTERVALS USING FORECASTED INDEPENDENT VARIABLES

X1 FORECAST = (4000, 6000)

	PREDICTION INTERVAL LOW	HIGH
MODEL 1 (X1=4000)	0.65	0.80
MODEL 2 (X1=6000)	0.60	0.75
MODEL FORECAST	0.60	0.80

were also checked. The updated model was then used to generate forecasts of pilot retention rates for fiscal year 1988.

Summary

The procedures discussed in this chapter were used to develop, validate, and update the model. Several different models were generated in the development process. Criteria were then established to chose the best model, which was used in validation and updating. The results are presented in the next chapter.

IV. Findings and Analysis

Introduction

Several models were built and tested in the iterative development process. Independent variables were modified, added, and deleted from the models during this process. Remedial measures were implemented to correct for departures from the error term assumptions.

In this chapter, the results of the performance of the initial model are reported. Variance stabilizing techniques, used to correct for nonconstant error term variance of the initial model, are then discussed. Three models were developed that satisfied all the assumptions of the general linear model. The performance of these models are compared by applying decision criteria to choose the best model. The results of the validation tests on the best model are then presented. Finally, the model update results are discussed and forecasts of pilot voluntary retention rates are generated for 1988 using the best model.

Results of the Initial Model

The initial model, similar to the Econometric model used by analysts at the Pentagon, is defined by the following equation:

$$\begin{aligned} VOLRETj = \; & B0j + B1DUM7j + B2DUM8j + B3DUM9j + B4DUM10j + \\ & B5AIRj + B6ACOL7j + B7ACOL8j + B8ACOL9j + \\ & B11UNEMPj\text{-}1 \end{aligned} \qquad (7)$$

where

VOLRETj	=	voluntary pilot retention rate for a given year of service in a fiscal year j
DUMXj	=	indicator for X year of service (X yos intercept) in fiscal year j
AIRj	=	number of major, national, and turbojet airline hires in fiscal year j
ACOLXj	=	annualized cost of leaving the service for year of service group X in a fiscal year j
UNEMPj-1	=	the unemployment rate in fiscal year j - 1

The R2 and adjusted R2 were 0.928 0.910 respectively. Residual analysis showed that the residual variance decreased as the predicted values approached 1.0, as had been expected (Appendix A). A plot of residuals versus time showed possible serial correlation (Appendix B). Remedial measures to correct for these departures were then implemented.

Results of the Variance Stabilizing Techniques

Two statistical techniques were applied to the initial model to stabilize the variance of the error terms. These techniques were weighted least squares and transformation of the dependent variable (pilot voluntary retention rates). Each technique was applied separately to the model and the results were compared.

As a result of applying the weighted least squares technique to the model, the variance of the error terms was slightly more consistent over the range of the predictions. The variance of the residuals still seemed to decrease as the predictions approached 1.0. The residuals were plotted against the predicted values (see appendix C). The R2 and adjusted R2 dropped slightly from 0.928 0.910 to 0.900 0.870.

The parameter estimates did not differ significantly from the unweighted model. Table II shows the regression coefficients before and after applying the weighted least squares technique.

TABLE II. RESULTS OF THE WEIGHTED LEAST SQUARES TECHNIQUE

INDEPENDENT VARIABLES	PARAMETER ESTIMATES	
	WITHOUT WLS	WITH WLS
INTERCEPT	0.750	0.771
DUM7	-0.434	-0.424
DUM8	-0.288	-0.263
DUM9	-0.293	-0.241
DUM10	-0.024	-0.019
AIR	-0.038	-0.036
ACOL7	0.039	0.039
ACOL8	0.025	0.023
ACOL9	0.028	0.024
UNEMP	0.032	0.029

The second variance stabilizing technique implemented was dependent variable transformation. An upper bound logarithmic transformation was performed on the pilot retention rates. The transformed variable, defined in equation 5, was designed to provide more constant error term variance over the range of predicted values. This transformation technique made a noticeable improvement in stabilizing the variance of the error terms (Appendix D). The R2 and adjusted R2 increased to 0.937 0.921, but the ACOL measures for years of service 7 and 8 were no longer significant (with p-values of 0.581 and 0.288 respectively). Thus, there was a need to find another predictor of the relative wage differences between the military and the civilian labor force. Also, because the error terms appeared

correlated, additional economic indicators were considered for use in the model.

Results of Models with New Predictors

The pay compensation indicator was added to the model as a possible replacement predictor for the ACOL measures. The pay compensation series was statistically significant (p-value of 0.0309) as a lagged variable. This new model, which includes airline hires, pay compensation lagged one year, and the unemployment rate lagged one year, as predictors, will be referred to as the pay model. The logarithmic transformation of the pilot retention rate is the dependent variable. The R2 and adjusted R2 are 0.933 0.920. The assumptions of normality, independence, and constant variance of the error terms are satisfied. The log transformation eliminated the variance bias due to the upper bound constraint (1.0) in the retention rates. The addition of pay compensation as a predictor eliminated the serial correlation of the error terms. The plot of the residuals against time showed no distinguishable pattern (Appendix E). The Durbin-Watson statistic was close to two (2.064), also indicating little chance of serial correlation.

The independent variables were tested for the presence of multicollinearity, the correlation of independent variables among themselves. The Variance Inflation Factors (VIF), described in chapter 3, were used to detect the presence of multicollinearity. VIF values greater than ten

are an indication of possible multicollinearity. All of the independent variables in this model have VIF values of less than two. Because two of the three independent variables are lagged, only one (airline hires) needs to be estimated in order to predict the pilot retention rates.

Two additional economic indicators were added to determine if they could significantly improve the prediction capability of the model. Each variable was added separately to determine its individual contribution to the model.

The index of help-wanted advertising in newspapers was first added to the pay model. This variable was significant only as an unlagged series. By adding this index to the model, pay compensation was no longer significant. The significant predictors in this model, referred to as the job model, are airline hires, help-wanted advertising in newspapers, and the unemployment rate lagged one year. The model has an R2 and adjusted R2 of 0.948 and 0.938, and the error term assumptions are satisfied (Appendix F). The prediction potential is reduced, however, because two variables are unlagged and have to be estimated when predicting pilot voluntary retention rates.

The corporate profits variable was then added to the pay model. This variable was significant only as a lagged series, which is intuitively appealing because it is a leading economic indicator. The pay compensation variable was not significant when corporate profits was included as a predictor. The error term assumptions are satisfied

(Appendix G) and the R2 and adjusted R2 are 0.946 and 0.936. Two of the three predictors are lagged. This model, referred to as the profit model, includes the following predictors: airline hires, corporate profits lagged one year, and the unemployment rate lagged one year.

Choosing the Best Model

The three models which meet the assumptions of the general linear model are the pay, job, and profit models. Three criteria were established to help choose the best model. These criteria are model fit, prediction potential, and explanatory significance. For model fit the following diagnostics were compared: the R2 and adjusted R2 values, the width of the prediction intervals, and the values of the mean square error. Because unlagged variables have to be estimated when forecasting pilot voluntary retention rates, the second criterion is that the best model should have the least number of unlagged variables. The final criterion is a comparison of the models based on the number of predictor types represented. The three types of predictors, described in the data description section, are indicators of the airline industry, the relative wage differences, and the economic indicators.

The three models were first rated based on their performance according to the model fit criterion. All three models provide good fit to the pilot retention data. Each model has an R2 above 0.90, which can be considered very

good. The prediction interval widths and mean square errors are similar for all three models (Table III).

TABLE III. MODEL FIT RESULTS

MODEL	R2	ADJ R2	MSE	INTERVAL WIDTH*
PAY	0.933	0.920	0.301	0.0508
JOB	0.948	0.938	0.232	0.0436
PROFIT	0.946	0.936	0.241	0.0445

* the 95% prediction interval widths are calculated based on the average interval for the five years of service for fiscal year 1985

The table values demonstrate that a distinction between models is difficult based solely on the model fit criteria.

However, one model can be eliminated from further consideration based on the second criterion. The job model has two unlagged series, which requires forecasts of two variables to be useful. The other two models require forecasts only of one variable, airline hires. Therefore, the job model has greater potential for prediction error and is considered to be the worst of the three in terms of a prediction model.

The third criterion helps distinguish the best model from the two remaining candidates. The pay model with independent variables airline hires, the unemployment rate, and pay compensation, has predictors of all three types (airline industry, economic indicators, and relative wage differences). The profit model, which has corporate profits,

the unemployment rate, and airline hires as independent variables, does not have a predictor for the relative wage difference between the military and the civilian labor force.

Based on the three criteria discussed, the pay model was selected as the best overall model. The profit model, which is equal to the pay model in the number of lagged variables, is suggested as an alternative model. These two models are updated with the 1986 and 1987 data.

Validating the Model

The actual pilot voluntary retention rates for fiscal year 1986 were compared to the predicted pilot voluntary retention rates from the pay model. That model is defined by the following equation:

$$\text{TRANSRET}j = \text{B0}j + \text{B1DUM7}j + \text{B2DUM8}j + \text{B3DUM9}j + \text{B4DUM10} + \text{B5AIR}j + \text{B6PAY}j\text{-1} + \text{B7UNEMP}j\text{-1} \qquad (8)$$

where

TRANSRETj	=	- ln(1.0 - voluntary retention + 0.01) in year j
DUMXj	=	X year of service indicator (for X yos intercept) in year j
AIRj	=	number of major, national, and turbojet hires in year j
PAYj-1	=	pay compensation in year j - 1
UNEMPj-1	=	unemployment rate in year j - 1

This model was used to generate prediction estimates and prediction intervals for the retention rates for forecasts one year ahead. A level of significance of 0.05 was used in computing the prediction interval widths.

The only unknown predictor variable is the number of airline hires. One year forecasts of this variable are needed to predict the pilot retention rates. The forecasts of the airline hires were provided by the USAF Retention Division at HQ AFMPC. The analysts at the Retention Division obtained the estimates by carefully studying the growth of the airline industry. The airline hire forecast for 1986 was treated as an interval estimate. The lower and upper bounds of the estimate were used to generate the prediction interval forecasts of the pilot retention rates by the method described and illustrated in chapter 3. The results of the above procedure are listed in Table IV.

TABLE IV. PAY MODEL FORECASTS FOR 1986

YEAR OF SERVICE	ACTUAL	PREDICTED	PREDICTION INTERVAL LOWER	UPPER
7	0.724	0.627	0.402	0.769
8	0.792	0.762	0.617	0.854
9	0.826	0.828	0.722	0.896
10	0.848	0.878	0.800	0.927
11	0.876	0.909	0.850	0.947

Two variants of the pay model were constructed to determine whether the width of the prediction interval could be decreased while maintaining the same level of significance. Because a forecast of airline hires was necessary, the width of the prediction intervals are larger

than the intervals would be with a model of similar fit and all lagged predictors. The two variants are the pay model with airline hires lagged one year, and the pay model without airline hires as a predictor. For the model with airline hires lagged one year, all independent variables are significant (for a 0.05 level of significance), and the general linear model assumptions are satisfied. But the R2 and adjusted R2 dropped to 0.866 and 0.841 respectively. As a result, the prediction interval widths are larger than those of the original pay model. The larger interval widths can be attributed to a larger mean square error, primarily because the unlagged series of airline hires is a better predictor than the lagged series.

The pay model without airline hires as a predictor has an R2 and adjusted R2 of 0.841 and 0.816. The interval widths are larger than those of the pay model, because the mean square error is larger. Table V summarizes the comparisons of the three models.

TABLE V. RESULTS OF THE PAY MODEL WITH DIFFERENT FORMS OF THE AIRLINE HIRES PREDICTOR

MODEL	R2	ADJ R2	INTERVAL WIDTH*
PAY	0.933	0.920	1.11
PAY WITH AIRLINE HIRES LAGGED	0.866	0.841	1.19
PAY WITHOUT AIRLINE HIRES	0.841	0.816	1.26

* based on the average of the interval widths for the five year of service groups for 1986 (transformed pilot retention rates)

The 1987 data were available in time to perform an additional validation test on the pay model. This model was first updated with the 1986 data. The results of this update (Appendix H) showed no departures from the error term assumptions, and the parameter estimates did not change greatly. The forecast for airline hires, used for these 1987 predictions, was given as a point estimate. So, the prediction intervals generated from this model are expected to cover the actual pilot voluntary retention rates with a probability of slightly less than 0.90. Table VI includes a comparison of the actual and predicted pilot voluntary retention rates.

TABLE VI. PAY MODEL FORECASTS FOR 1987

YEAR OF SERVICE	ACTUAL	PREDICTED	PREDICTION INTERVAL LOWER	UPPER
7	0.606	0.582	0.387	0.716
8	0.724	0.728	0.600	0.817
9	0.787	0.801	0.705	0.866
10	0.859	0.855	0.784	0.903
11	0.856	0.891	0.836	0.928

Model Updates

Both the pay and the profit models were updated to include data from 1977 through 1987. Each updated model was analyzed for model fit, departures from the error term assumptions, and changes to the parameter estimates. The

parameter estimates of the two updated models were compared to the estimates of their original versions.

The additional data made small changes to the R2, adjusted R2, and the parameter estimates of the pay model. Residual analysis suggested no departures from the general linear model (Appendix I). The results are summarized in Table VII.

TABLE VII. PAY MODEL UPDATE RESULTS

	1985 MODEL	1987 MODEL
R2	0.933	0.925
ADJUSTED R2	0.920	0.914
INTERVAL WIDTH*	0.097	0.111
PARAMETER ESTIMATES		
INTERCEPT	-9.26	-8.41
DUM7	-3.83	-3.44
DUM8	-2.58	-2.30
DUM9	-1.69	-1.50
DUM10	-0.78	-0.65
AIRLINE	-0.78	-0.75
PAYCOMP	0.24	0.26
UNEMP	1.21	1.01

* based on one year ahead forecast of the 11 year of service group

The profit model was also updated because it is offered as an alternative model. The updated model, compared to the original model, has similar R2 and adjusted R2 values (Appendix J). The general linear model assumptions are satisfied, but some of the parameter estimates changed significantly. The coefficients for corporate profits and the unemployment rate differ significantly from those in the original profit model (Table VIII).

TABLE VIII. PROFIT MODEL UPDATE RESULTS

	1985 MODEL	1987 MODEL
R2	0.946	0.931
ADJUSTED R2	0.936	0.920
INTERVAL WIDTH*	0.094	0.101
PARAMETER ESTIMATES		
INTERCEPT	11.33	6.04
DUM7	-3.83	-3.44
DUM8	-2.58	-2.30
DUM9	-1.69	-1.50
DUM10	-0.78	-0.65
AIRLINE	-1.01	-1.16
PROFITS	-0.93	-0.45
UNEMP	0.45	0.79

* based on one year ahead forecast of the 11 year of service group

Because some of the parameter estimates significantly changed in the profit model, the pay model remains the preferred forecasting model.

Forecasts of Pilot Retention Rates for 1988

Using the updated 1987 pay model, forecasts of pilot voluntary retention rates for 1988 were generated. The 1988 airline hires forecast, provided by the USAF Retention Division, was treated as a point estimate. The prediction intervals of the pilot retention rates should cover the actual pilot retention rates with a probability of slightly less than 0.90. The results are summarized in Table IX.

Summary

Of the many models developed in this study, three actually demonstrated good model fit and satisfied the

TABLE IX. PAY MODEL FORECASTS FOR 1988

YEAR OF SERVICE	PREDICTED	PREDICTION INTERVAL LOWER	PREDICTION INTERVAL UPPER
7	0.4825	0.2546	0.6416
8	0.6606	0.5097	0.7660
9	0.7492	0.6365	0.8279
10	0.8184	0.7357	0.8762
11	0.8587	0.7934	0.9044

general linear model assumptions. These models were compared against the criteria of model fit, prediction potential, and explanatory significance. The pay model was selected as the best model.

Validation tests were performed on the pay model for 1986. This model was then updated and re-validated for 1987. Both the pay and profit models were updated through 1987. Finally, forecasts of the 1988 pilot voluntary retention rates were generated using the pay model.

The data used to generate and update these models are provided in Appendix K. With the aid of a regression software package, the models can be available for immediate use. Implications of the results of this chapter and recommendations for areas of further study are discussed in the final chapter.

V. *Conclusions and Implications*

Introduction

This chapter is a summary of the implications of the model results in performance and validation testing. The potential application of the model as a management tool will be addressed by identifying its strengths and limitations. In addition, several recommendations for refinements to the present model will be suggested as areas for further research.

Practical Implications of the Results

Using the results from the model development and validation tests, the researcher is able to assess the utility of the pilot retention model in terms of its value as a practical tool for analysts. The scope of this effort reflects the limitations of the application of the model. This effort focused on short term forecasts of pilot voluntary retention rates for seven through eleven years of service. The model's strengths are its prediction capability, and its simplicity. Also, the model shows that a statistical relationship exists between the retention rates and certain explanatory variables.

The results of the validation tests demonstrated the model's ability to predict pilot voluntary retention rates one year ahead. Each prediction interval covered the actual retention rate for each year group as a result of the two validation tests (for fiscal years 1986 and 1987). Although

the degree of prediction accuracy is subjective, the predictions from this model will help Air Force leadership better anticipate the voluntary separations of the pilots.

A statistical relationship exists between pilot retention rates and the predictors of the model. Pilot retention is statistically related to the number of airline hires, the unemployment rate, the pay compensation between the military and civilian labor force, and the profits of U.S. corporations. The existence of a relation between pilot retention rates and these series in a lag form demonstrates that some of these series are leading indicators of pilot retention.

The model is relatively simple and easy to maintain. Regression analysis is a common analytical tool and is known by many analysts. Each of the suggested models (the pay and profit models) have only three predictors. With the exception of the forecast for the number of airline hires, the data needed to update the models are readily available. The model can also be maintained and updated on a personal computer.

The assumptions made in defining the scope of the problem inherently limit the application of the model. The retention rates are only for the voluntary separations for year of service groups seven through eleven. The reasons for these limitations, discussed in the first chapter, are that pilots have service commitments until their seventh year and historically have remained in the service after their

eleven year (after promotion to Major). Pilot separations are a concern, mostly because of the time and cost involved in training new pilots. The voluntary separations are a concern because they are more numerous and variable than the involuntary separations. In addition, these forecasts are only for one year ahead. Forecasts beyond one year with this model would have large prediction interval widths, mostly because all the predictor variables would have to be estimated. The larger prediction interval width reduces the utility of the forecast.

Because only eleven years of data were available to build this model, annual updates are important. The updates ensure that the most recent information is used to build the model. The model updates require the data files to be modified to include the new data. Then, the regression analysis is performed to obtain the new parameter estimates for the predictions. The parameter estimates should be checked for large deviations from the estimates of the previous model. Also, residual analysis should be performed to ensure the assumptions of the general linear model are maintained.

Recommendations for Refinement

The suggestions for refining and improving the model begin with data problems. Data collection is the most important phase of the model development process. Obviously, with insufficient or inaccurate data, a researcher has problems generating a realistic model of significance. It is

important to ensure that the collected data is accurate before it is used in the model.

Another problem in data collection is finding data to support potential predictors. Some predictors may make sense intuitively, but the associated data does not exist. For instance, one possible indicator of the trend in retention in the short term would be direct survey questioning of pilot's intentions to separate. A pilot survey was conducted by the Officer Survey Branch at HQ AFMPC in January of 1987. However, the questions in the survey were not specific concerning separation. An annual survey of pilots without an Active Duty Service Commitment should be conducted to determine their separation intentions for the following year. This information would be extremely valuable to the analyst who is trying to predict this response.

The number of data points used to build the model could be greatly increased if quarterly data were available for each of the predictors, in addition to the pilot voluntary retention rates. By increasing the number of data points, a more accurate model could be built using regression analysis. Also, other techniques, such as time series analysis, could be implemented in an effort to improve the prediction capability of the model. If quarterly data were available for all the predictors, various other lagging schemes could be used.

The data for the number of new airline hires is possibly inflated recently due to pilots changing employers within the industry. Airlines are beginning to recruit pilots who are

actively flying for other airlines. As a result, the data collected does not reflect the actual new hires situation. An accurate count of the number of new hires of major, national, and turbojet companies is needed.

More accurate forecasts of the number of new airline hires are needed to refine the present model. The accuracy of these forecasts is directly related to the width of the prediction intervals. One possible approach is to build a general linear model using explanatory variables from the airline industry and the economy which are leading indicators of airline hires.

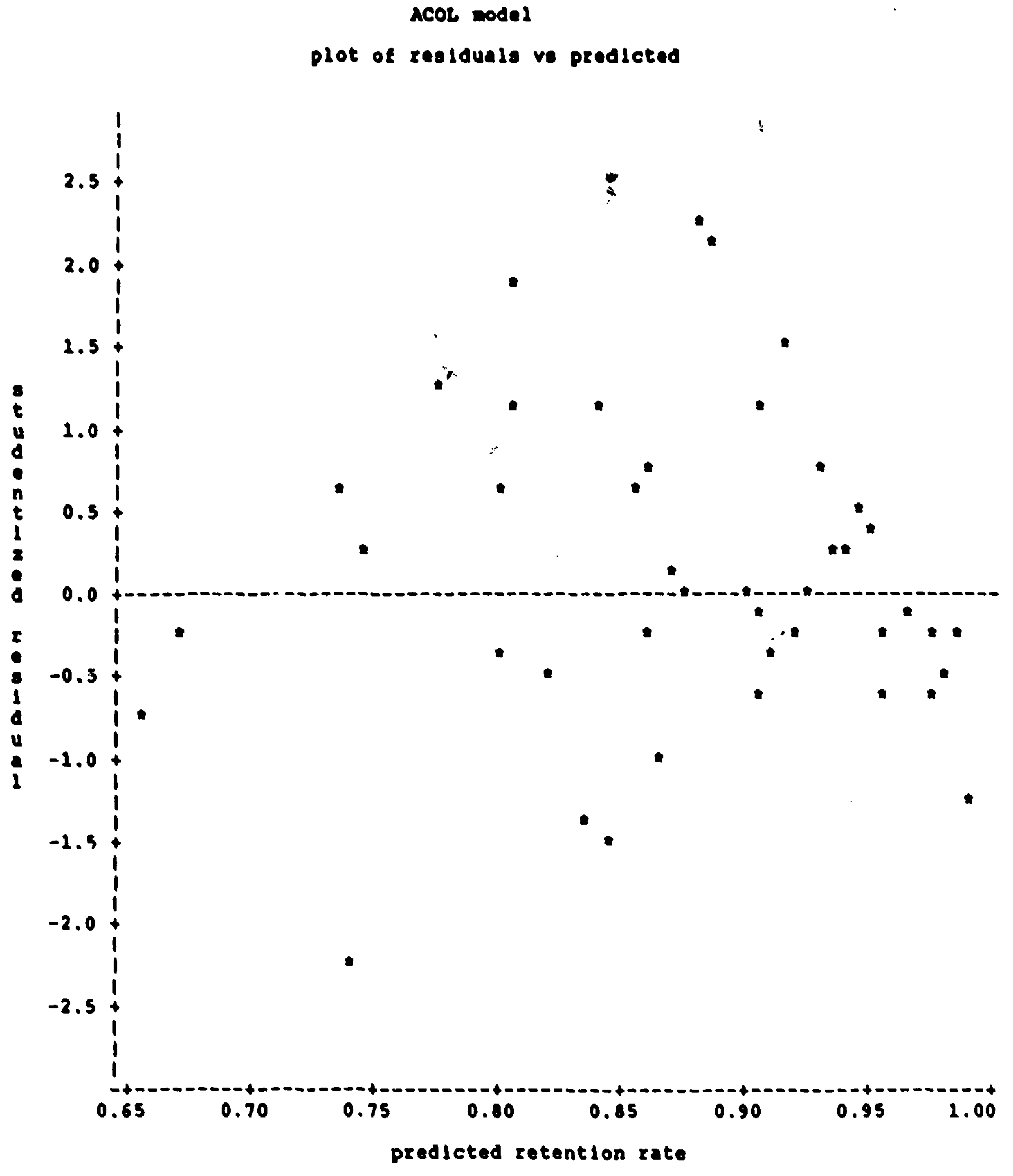
Another approach to refining the model developed in this study is to find other appropriate explanatory variables that are readily available and can improve the prediction capability of the model. Different economic indicators or other measures of the strength of the airline industry might improve the present model. An approach to account for the taste for military service, other than creating separate equations for each year of service group, might also improve the predictions.

Other enhancements to this model that would increase its utility include the following: 1) predict retention rates in the out years (2 or more years ahead), 2) predict involuntary retention rates for all year groups, 3) predict voluntary retention rates for year of service groups eleven through twenty eight, 4) predict retention rates by year of service group and by weapon system. These enhancements are

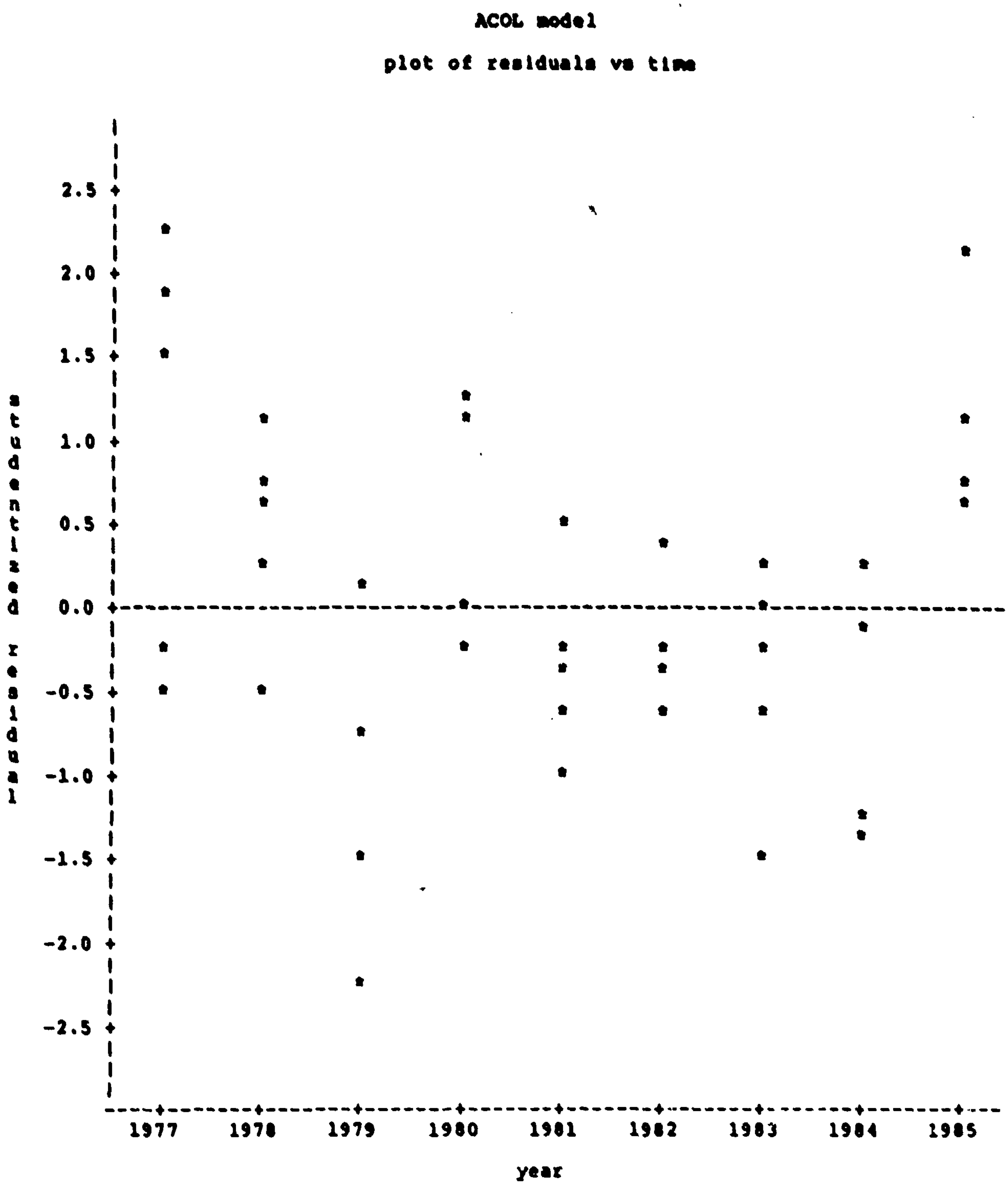
recommendations for areas of further research, and the best approach may be to develop a new model.

This research effort has produced a model which accurately forecasts pilot voluntary retention rates for year groups seven through eleven. Several explanatory variables have been shown to be statistically significant leading indicators of pilot retention. These findings will benefit those working in the area of pilot retention forecasts.

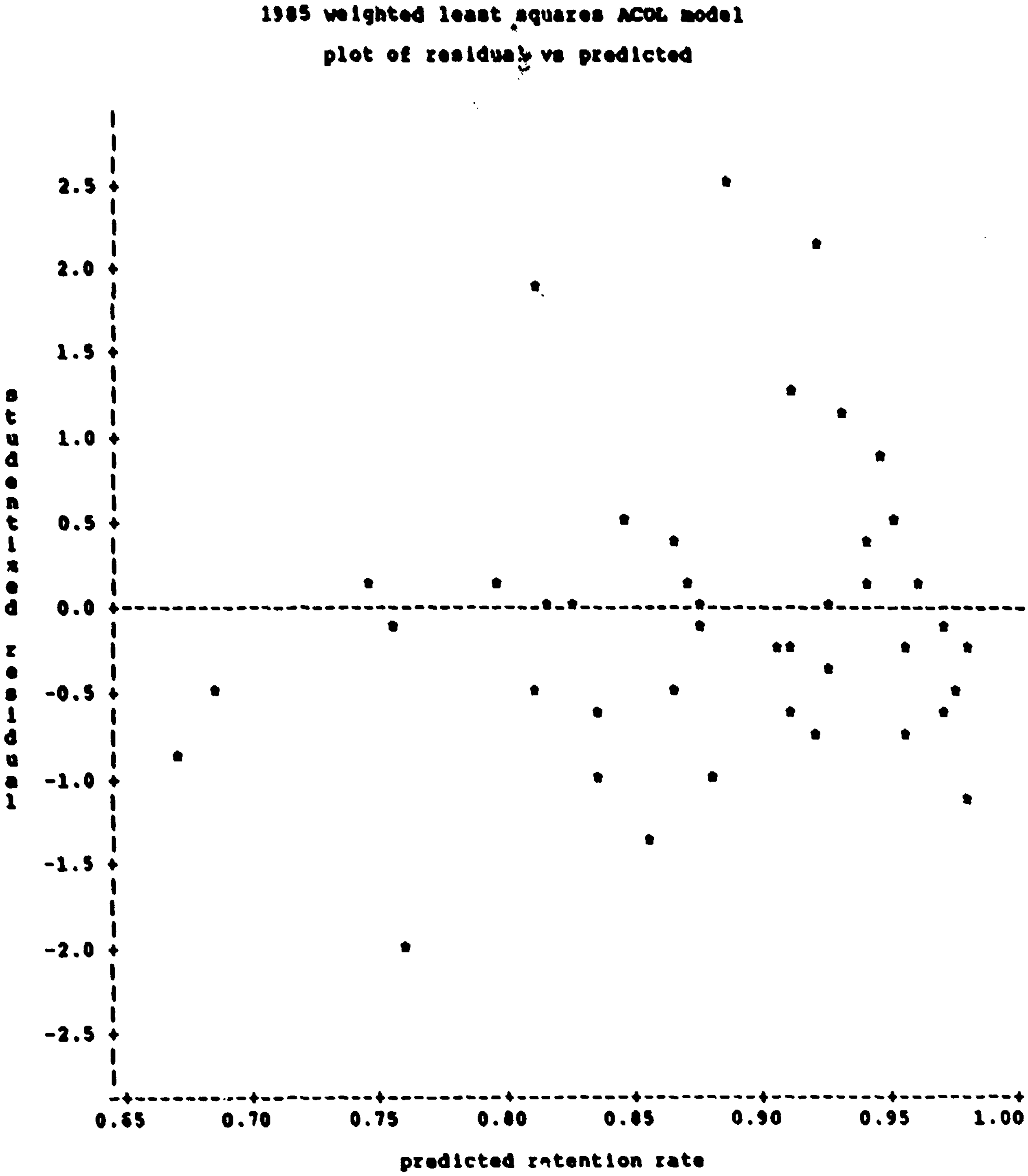
Appendix A: *Plot of Residuals versus Predicted for the ACOL Model*



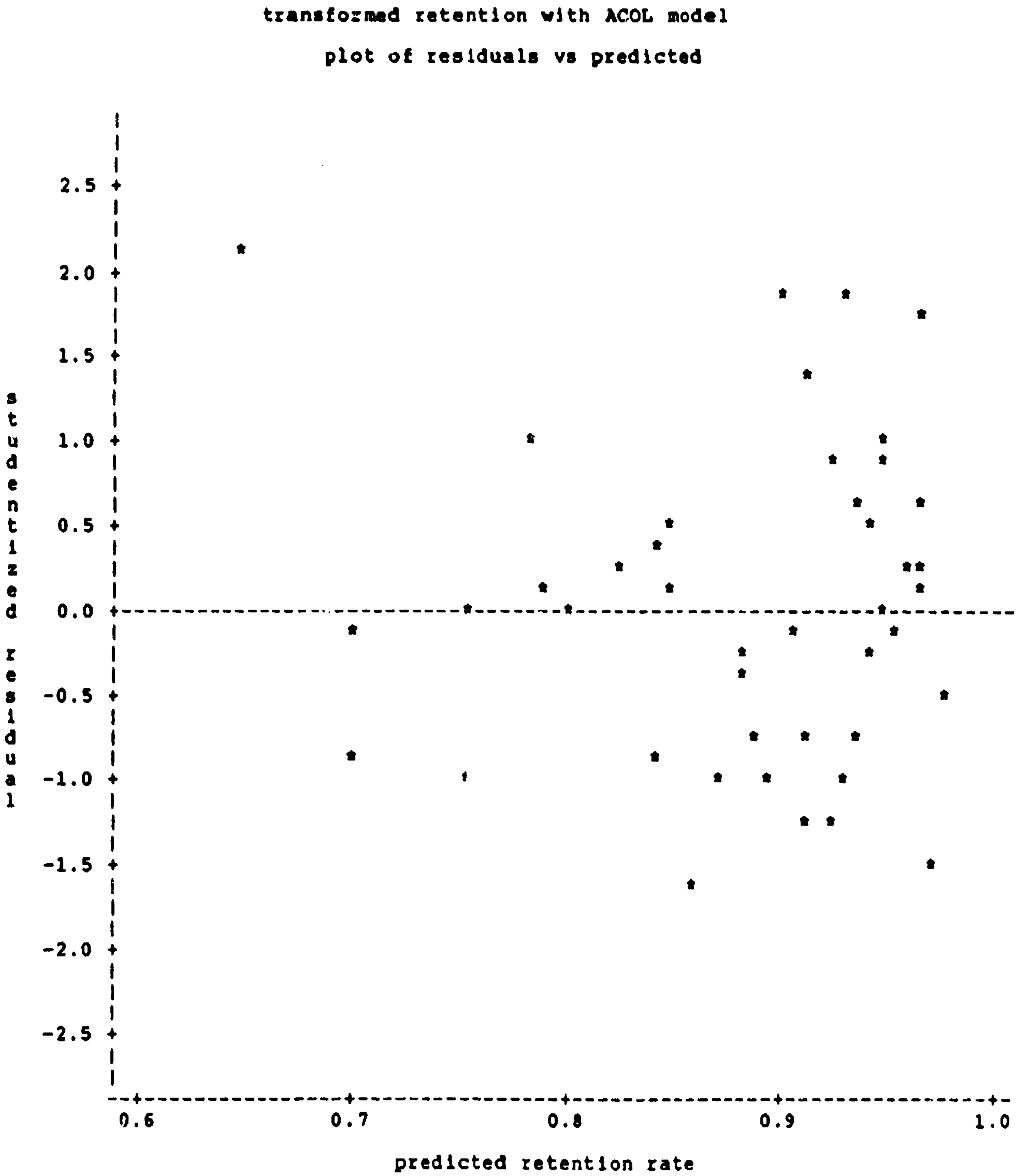
Appendix B: *Plot of Residuals versus Time for the ACOL Model*



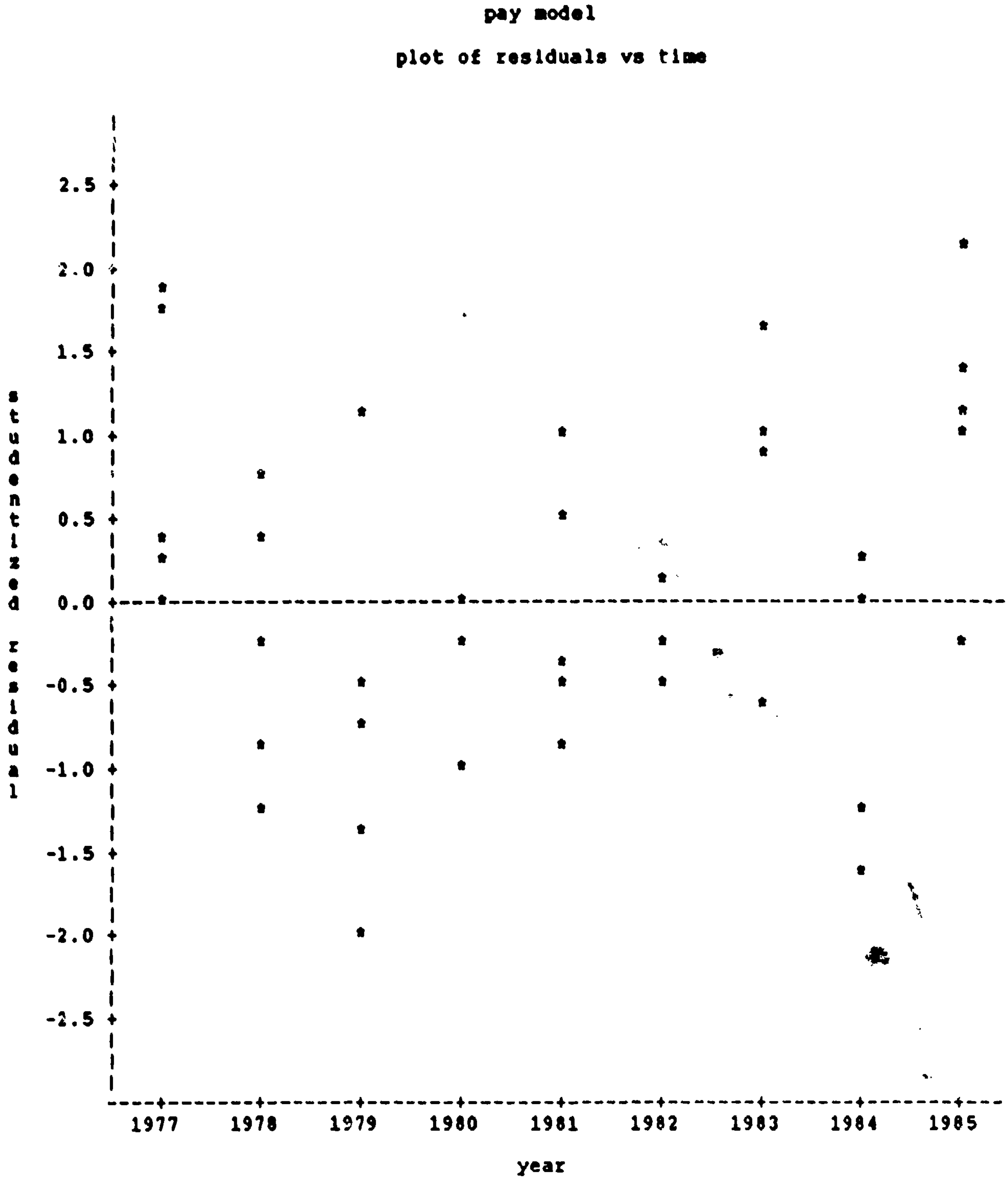
Appendix C: Plot of Residuals versus Predicted for the Weighted Least Squares Model



Appendix D: *Plot of Residuals versus Predicted for the ACOL Model with Transformed Retention Rates*

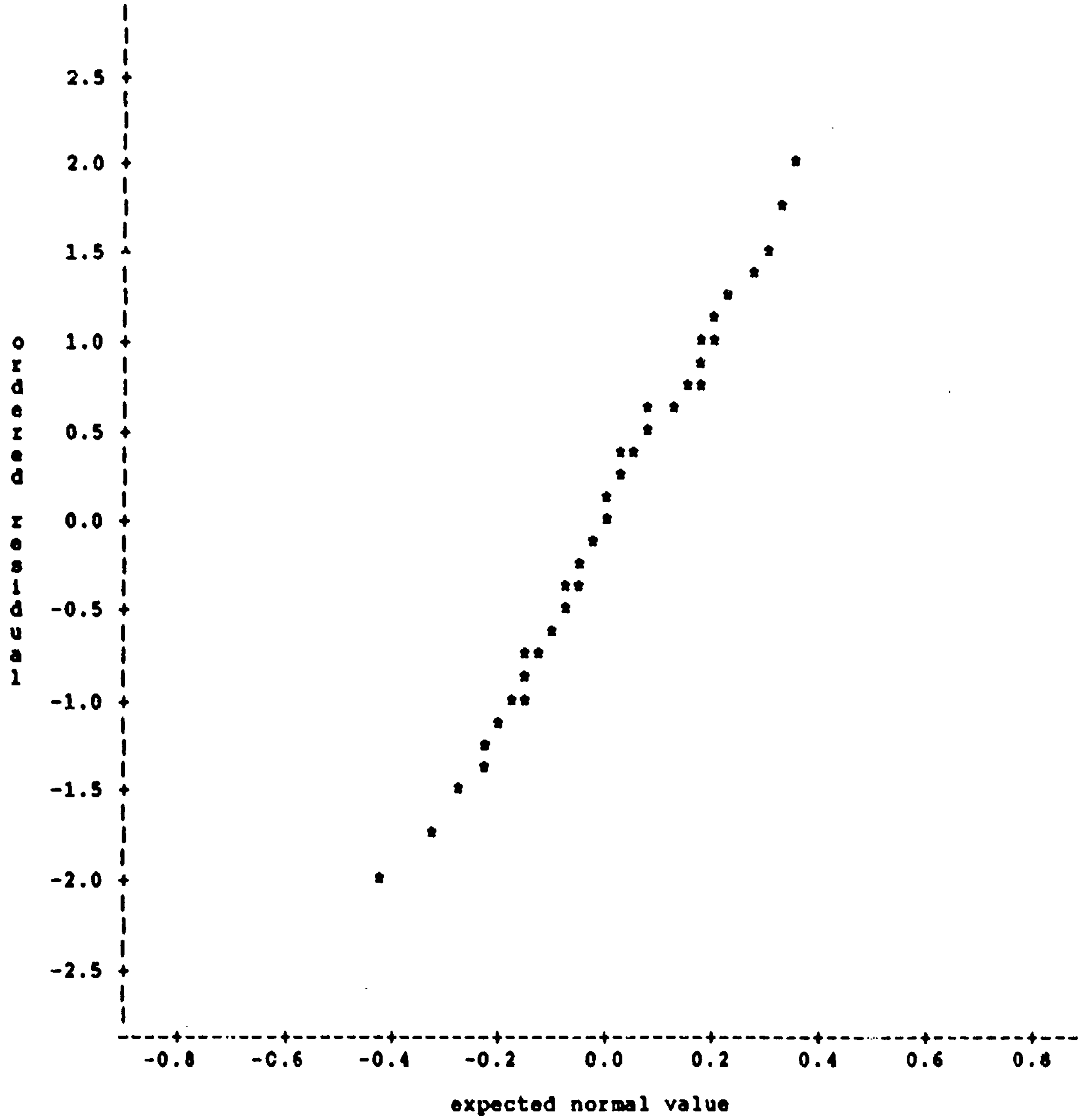


Appendix E: *Residual Analysis Plots and Analysis of Variance Table- Pay Model*



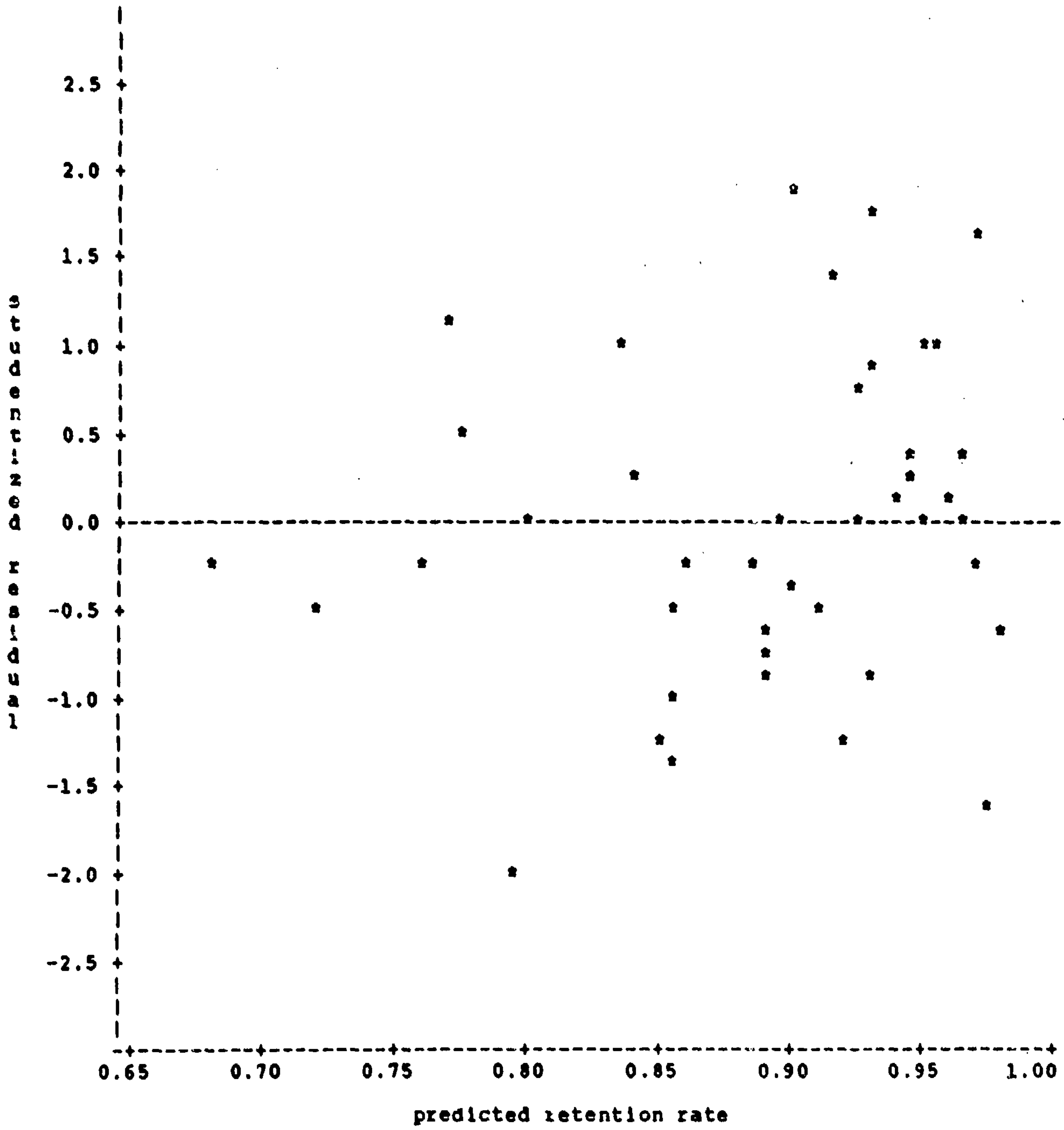
1985 pay model

normal probability plot



pay model

plot of residuals vs predicted



1935 pay model

ANALYSIS OF VARIANCE					
SOURCE	DF	SS	MS	F	PROB F
MODEL	7	154.44	22.06	73.4	0.0001
ERROR	37	11.12	0.30		
TOTAL	44	165.56			
R^2	0.933				
ADJ R^2	0.920				
DW STATISTIC	2.064				

PARAMETER ESTIMATES			
VARIABLE	ESTIMATE	T-STATISTIC	VIF
INTERCEPT	-9.26	-1.83	0
DUM7	-3.83	-14.83	1.6
DUM8	-2.58	-9.99	1.6
DUM9	-1.69	-6.56	1.6
DUM10	-0.77	-3.00	1.6
AIRLINE	-0.78	-7.09	1.6
PAY	0.24	2.24	1.5
Ur..	1.22	12.31	1.3

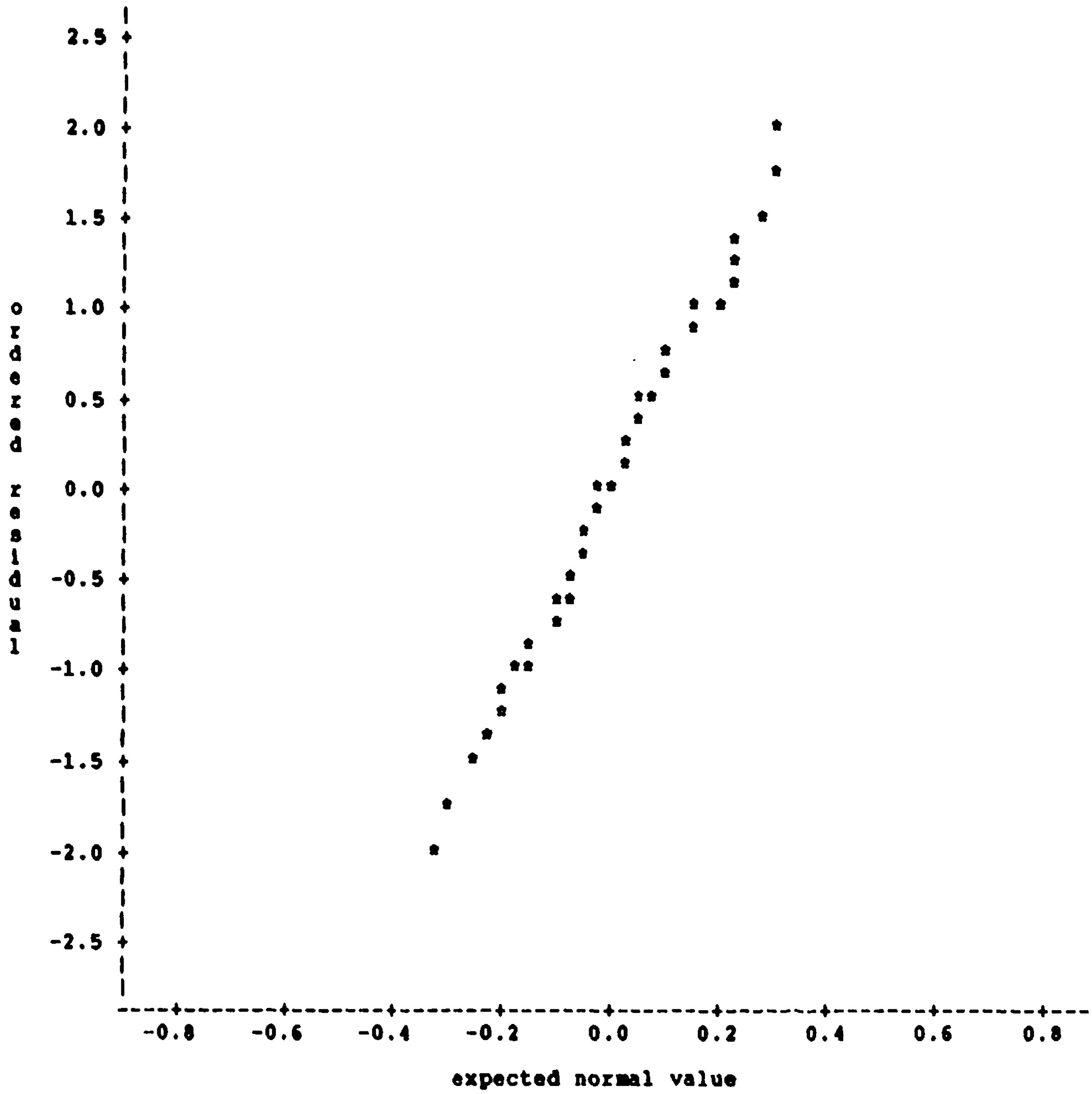
Appendix F: Analysis of Variance Table and Residual Analysis Plots- Job Model

ANALYSIS OF VARIANCE					
SOURCE	DF	SS	MS	F	PROB F
MODEL	7	156.966	22.424	96.5	0.0001
ERROR	37	8.598	0.232		
TOTAL	44	165.563			
R^2	0.948				
ADJ R^2	0.938				
DW STATISTIC	1.780				

PARAMETER ESTIMATES			
VARIABLE	ESTIMATE	T-STATISTIC	VIF
INTERCEPT	7.36	5.45	0
DUM7	-3.83	-16.87	1 6
DUM8	-2.58	-11 37	1 6
DUM9	-1.69	-7 45	1.6
DUM10	-0.77	-3.41	1 6
AIRLINE	-0.45	-3 24	3 3
JOB	-0.63	-4.17	3 9
UNEMP	0.85	6 23	3 2

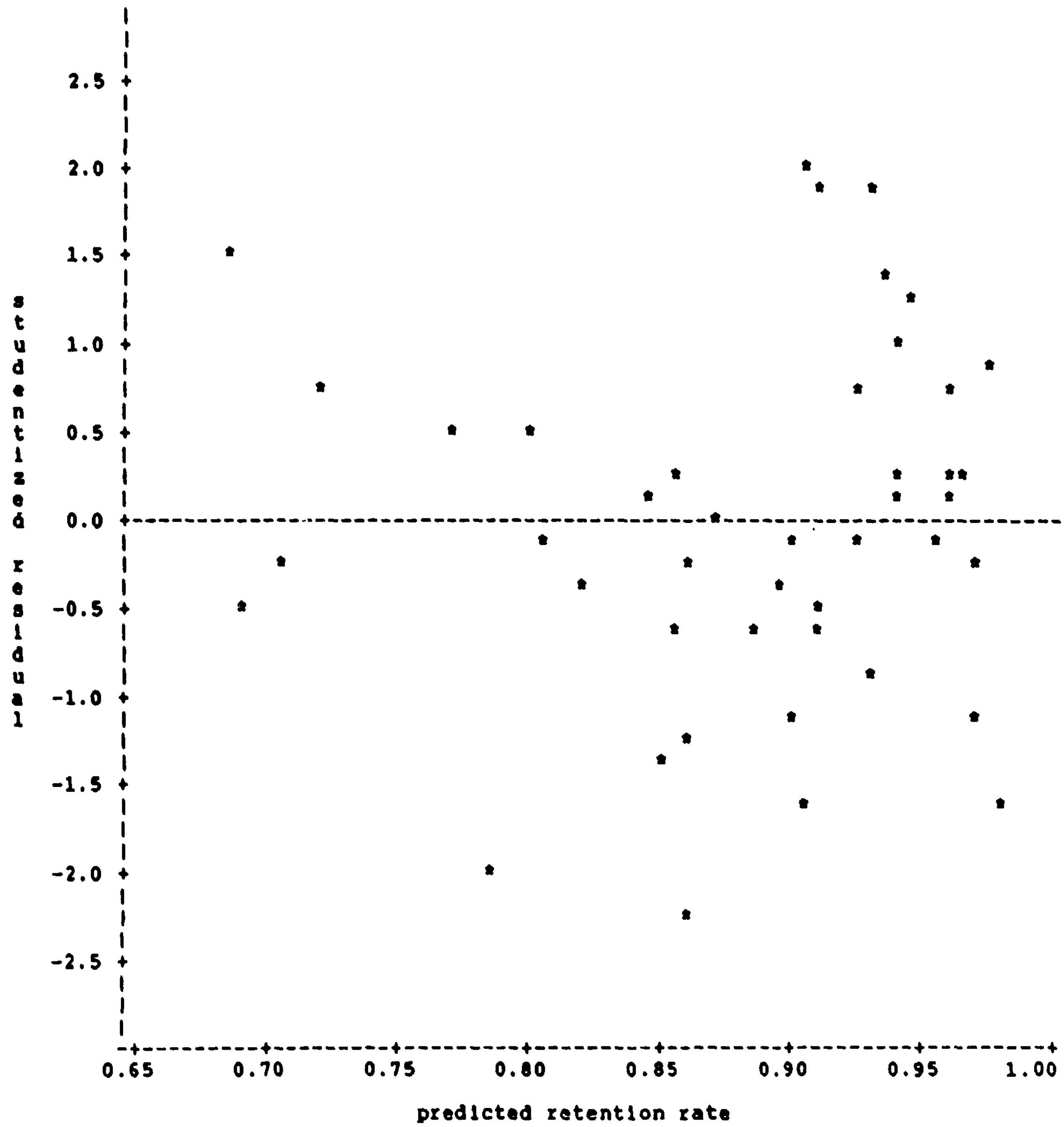
1985 job model

normal probability plot



job model

plot of residuals vs predicted



Appendix G: Analysis of Variance Table and Residual Analysis Plots- Profit Model

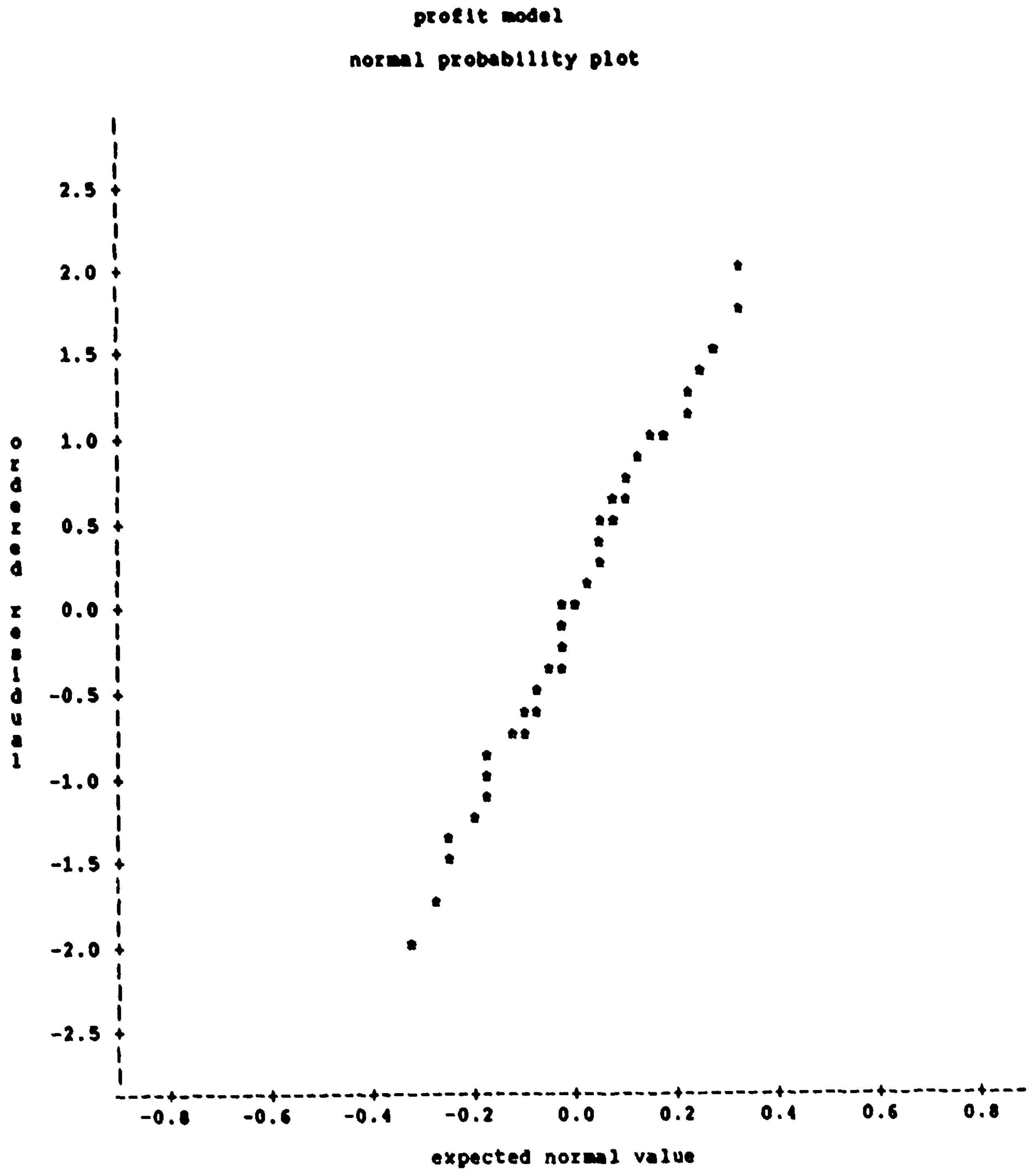
ANALYSIS OF VARIANCE

SOURCE	DF	SS	MS	F	PROB F
MODEL	7	156.64	22.38	92.8	0.0001
ERROR	37	8.92	0.24		
TOTAL	44	165.56			

R^2	0.946
ADJ R^2	0.936
DW STATISTIC	1.731

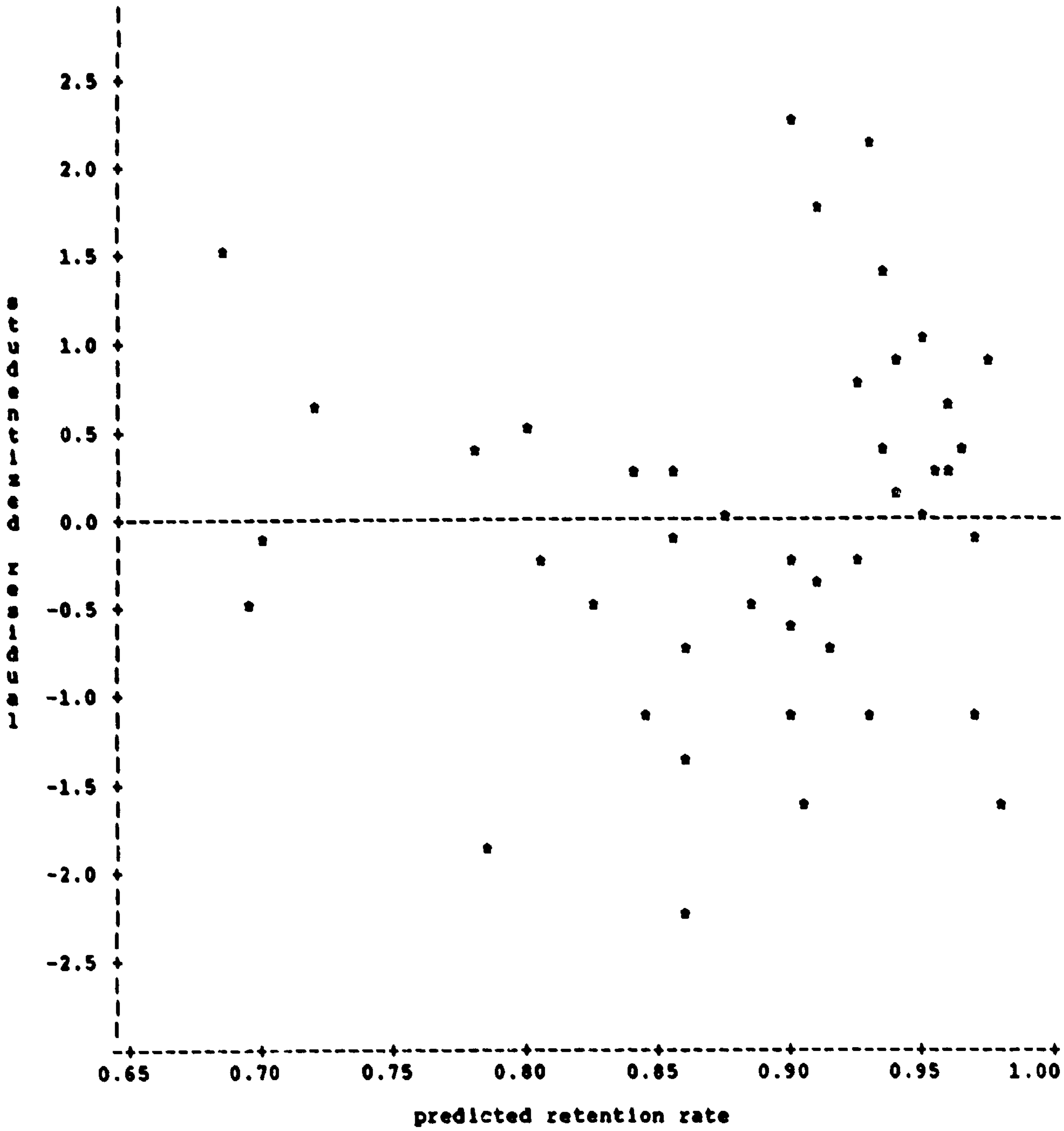
PARAMETER ESTIMATES

VARIABLE	ESTIMATE	T-STATISTIC	VIF
INTERCEPT	11.33	4.71	0
DUM7	-3.83	-16.57	1.6
DUM8	-2.58	-11.17	1.6
DUM9	-1.69	-7.32	1.6
DUM10	-0.77	-3.35	1.6
AIRLINE	-1.01	-11.94	1.2
PROFIT	-0.93	-3.93	9.4
UNEMP	0.45	1.97	8.8


profit model
normal probability plot
ordered residual
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-0.8
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
0.8
expected normal value

profit model

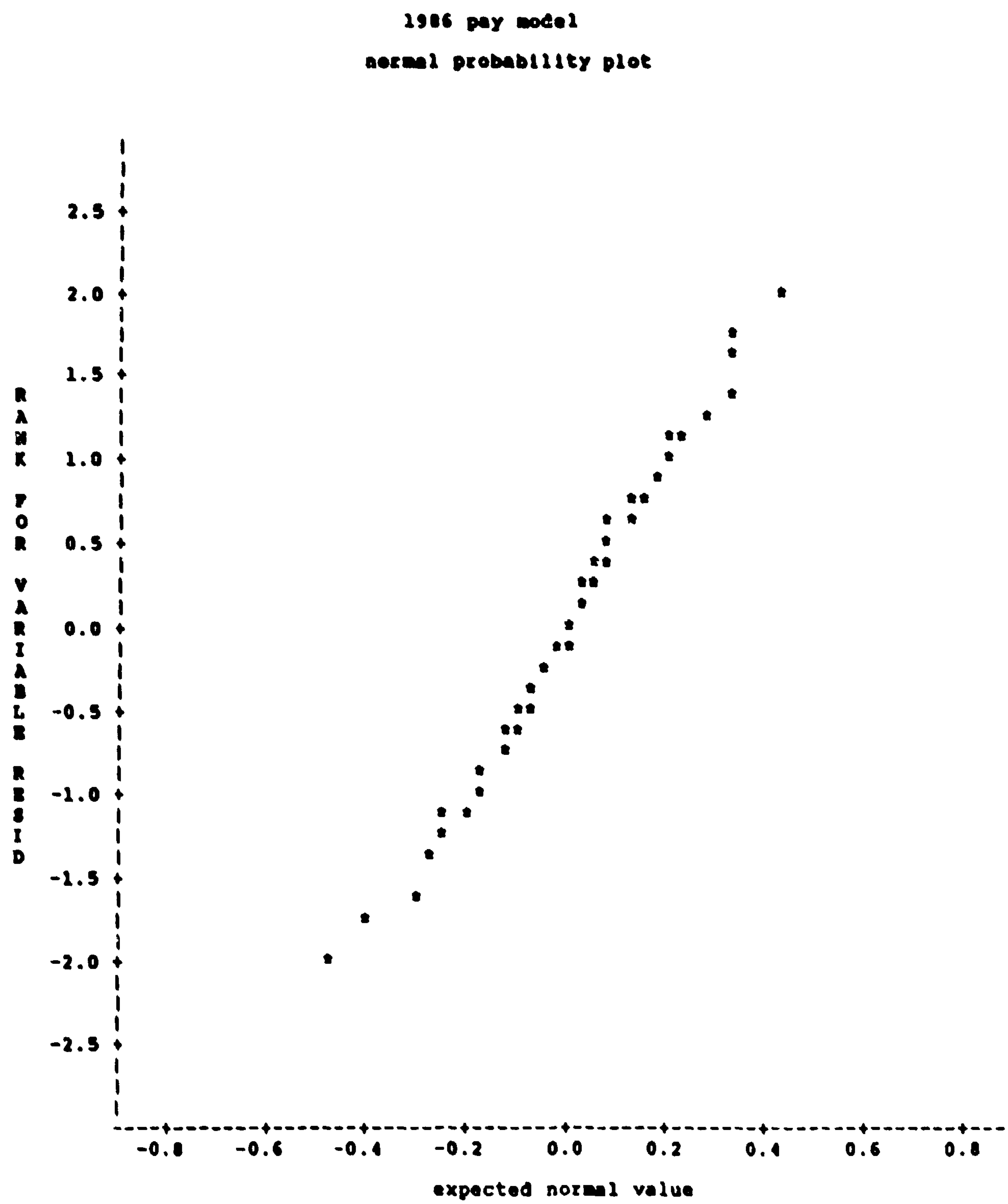
plot of residuals vs predicted



Appendix H: Analysis of Variance Table and Residual Analysis Plots- 1986 Pay Model

ANALYSIS OF VARIANCE					
SOURCE	DF	SS	MS	F	PROB F
MODEL	7	175 73	25.10	67 3	0 0001
ERROR	24	15 67	0 37		
TOTAL	49	191 39			
R^2	0.918				
ADJ R^2	0.905				
DW STATISTIC	1 731				

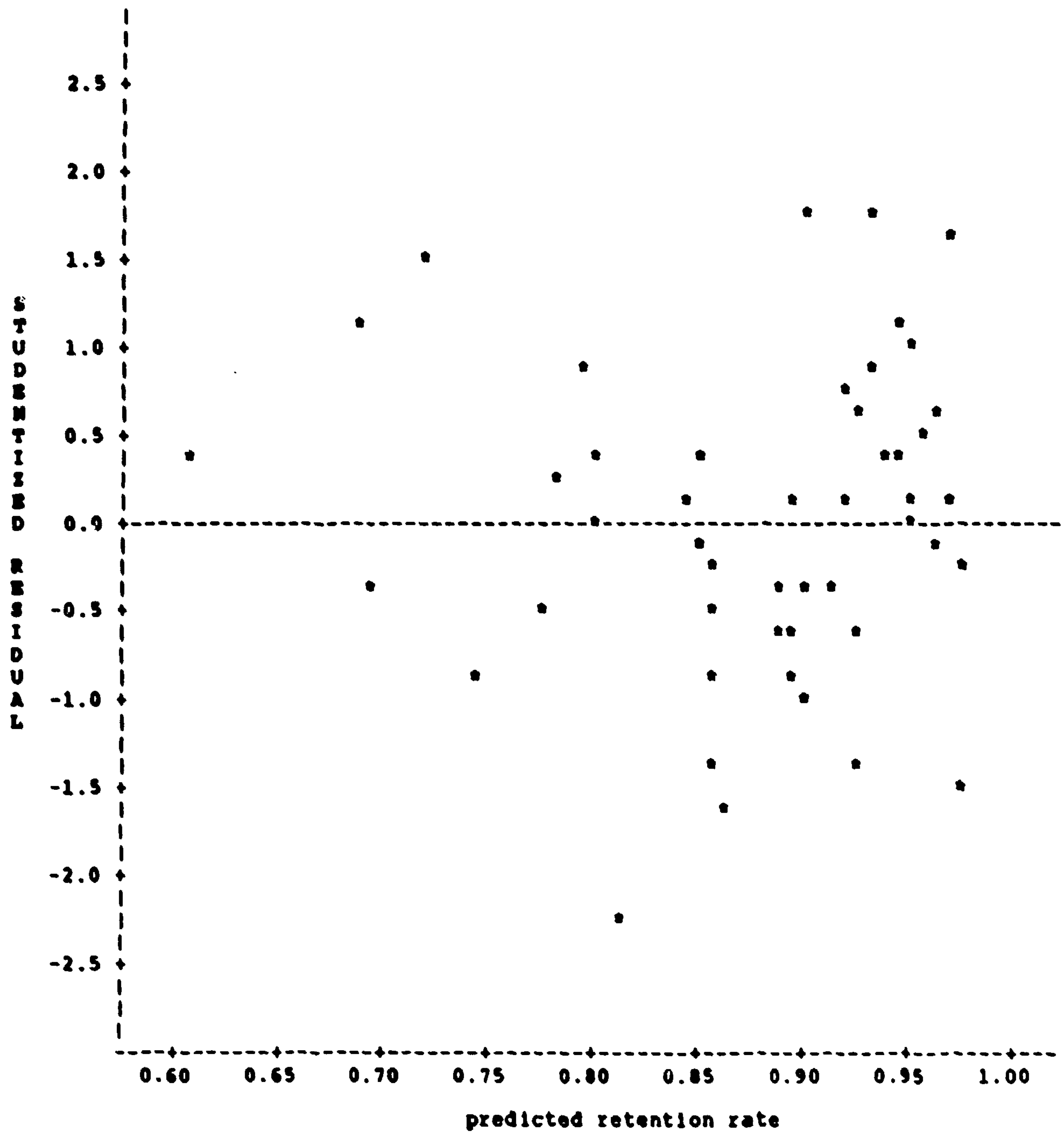
PARAMETER ESTIMATES			
VARIABLE	ESTIMATE	T-STATISTIC	VIF
INTERCEPT	-8.94	-1.58	0
DUM7	-3.78	-13.86	1.6
DUM8	-2.54	-9 32	1.6
DUM9	-1.67	-6.10	1 6
DUM10	-0 78	-2 85	1 6
AIRLINE	-0.79	-6 23	1 9
PAY	0.27	1 99	1 9
UNEMP	1 08	11.16	1 2



1986 pay model
normal probability plot
RANK FOR VARIABLE RESID
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-0.8
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
0.8
expected normal value

1986 pay model

plot of residuals vs predicted



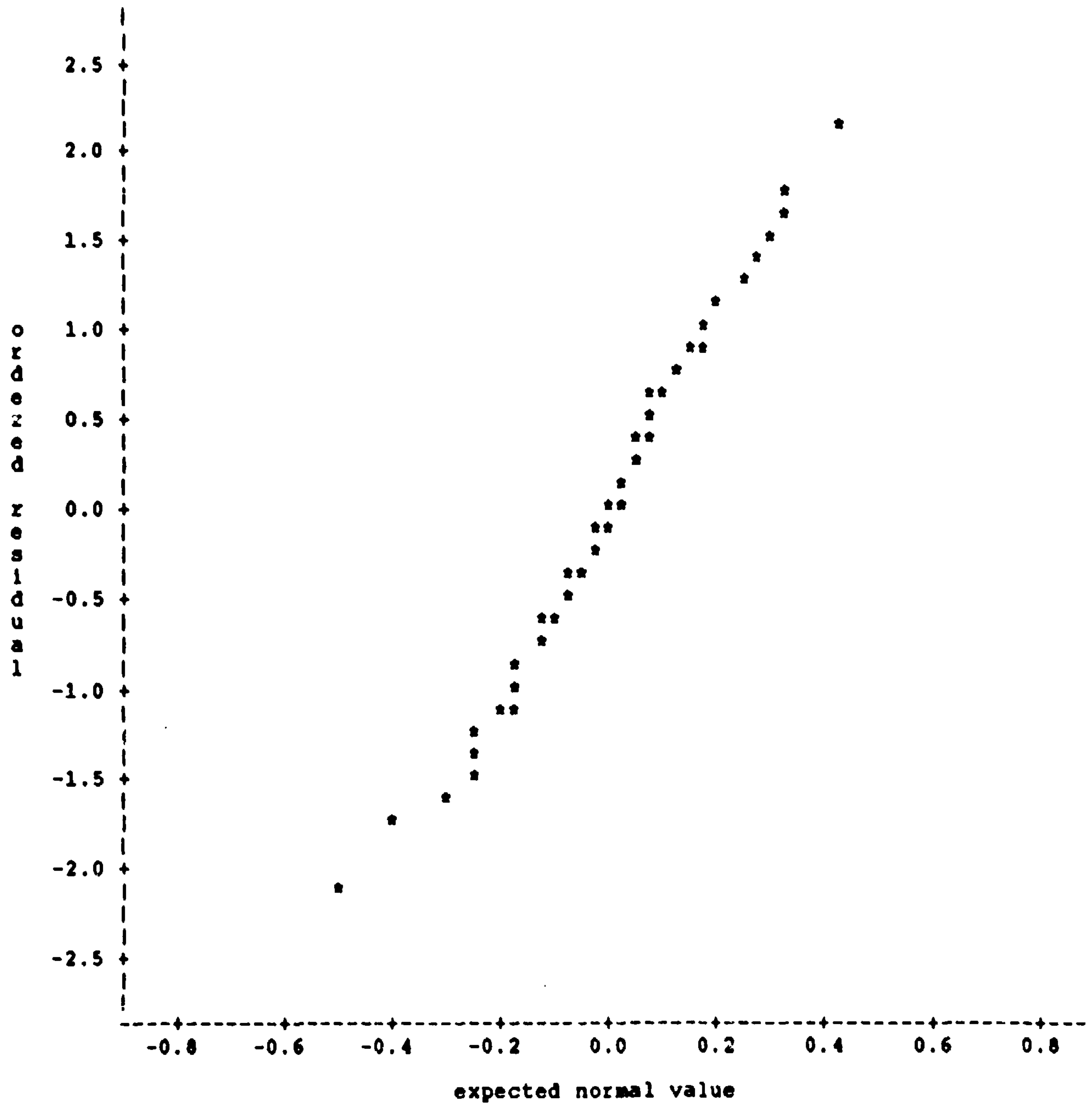
Appendix I: *Analysis of Variance Table and Residual Analysis Plots- 1987 Pay Model*

ANALYSIS OF VARIANCE					
SOURCE	DF	SS	MS	F	PROB F
MODEL	7	172.20	24.60	82.96	0.0001
ERROR	47	13.94	0.30		
TOTAL	54	186.14			
R^2	0.925				
ADJ R^2	0.914				
DW STATISTIC	1.867				

PARAMETER ESTIMATES			
VARIABLE	ESTIMATE	T-STATISTIC	VIF
INTERCEPT	-8.41	-1.76	0
DUM7	-3.44	-14.80	1.6
DUM8	-2.30	-9.92	1.6
DUM9	-1.49	-6.45	1.6
DUM10	-0.65	-2.80	1.6
AIRLINE	-0.75	-6.53	2.2
PAY	0.26	2.19	2.3
UNEMP	1.01	11.65	1.3

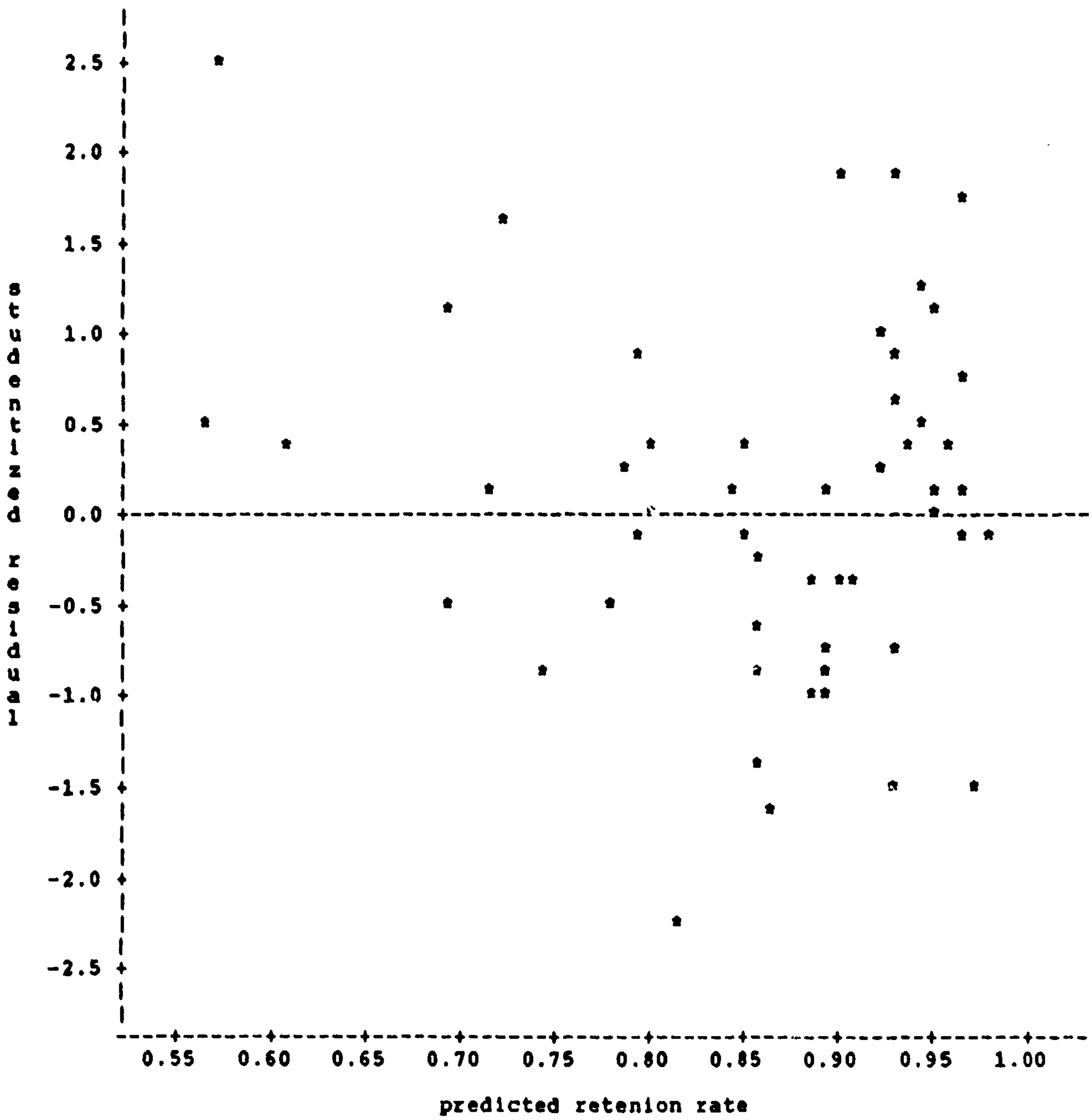
1987 pay model

normal probabilty plot



1987 pay model

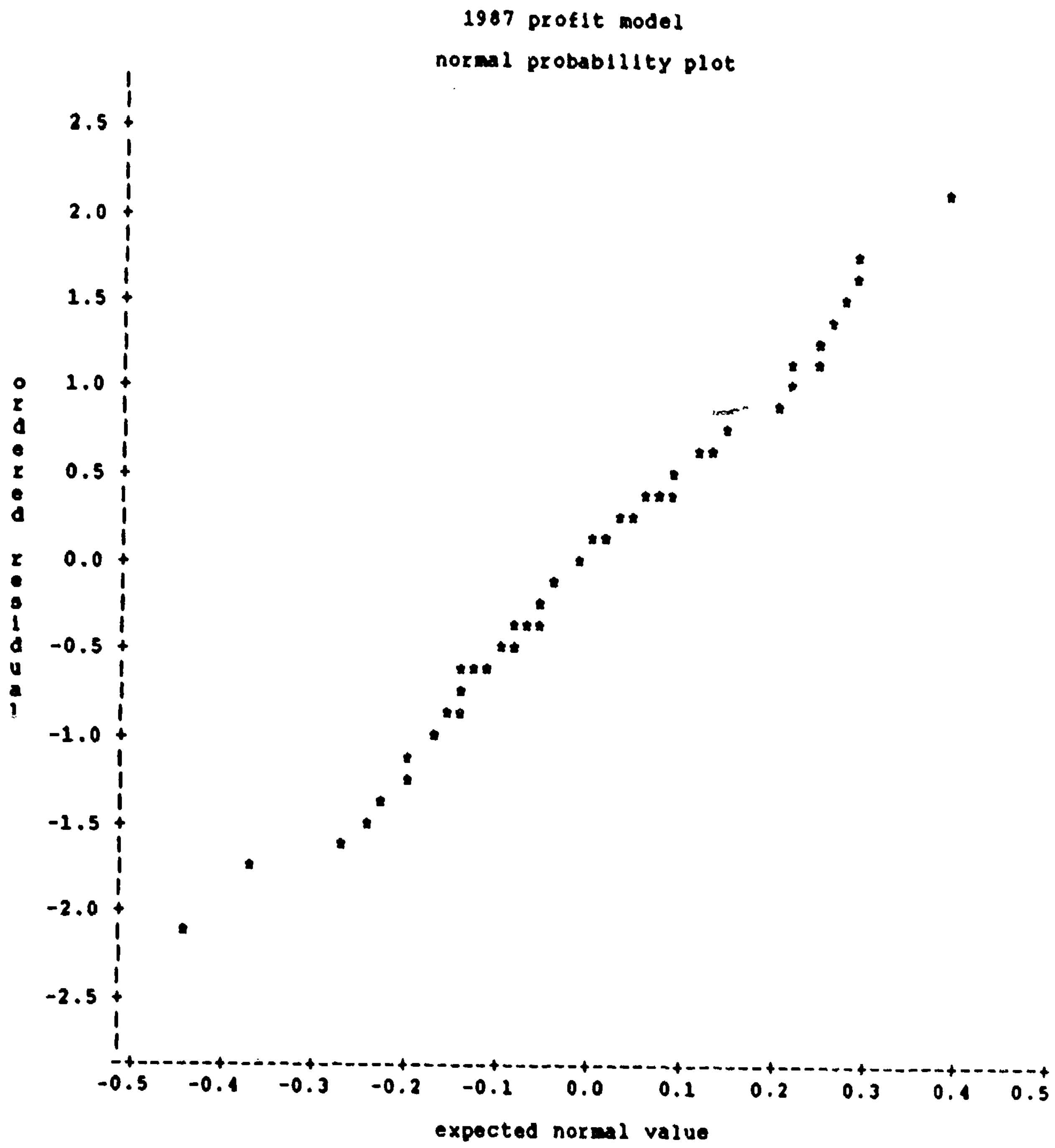
plot of residuals vs predicted



Appendix J: Analysis of Variance Table and Residual Analysis Plots- 1987 Profit Model

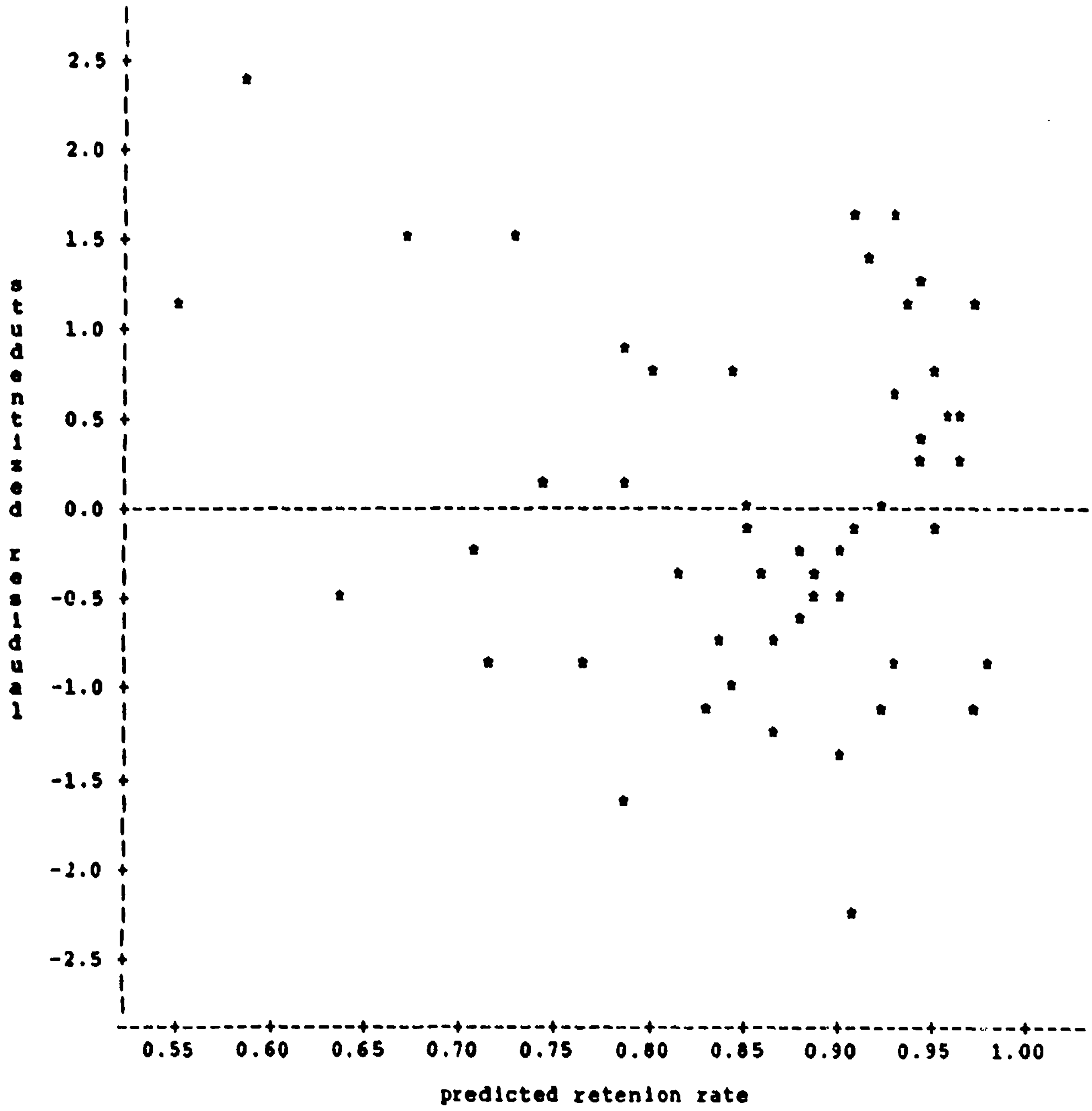
ANALYSIS OF VARIANCE					
SOURCE	DF	SS	MS	F	PROB F
MODEL	7	173.21	24.74	89.9	0.0001
ERROR	47	12.93	0.28		
TOTAL	54	186.14			
R^2	0.931				
ADJ R^2	0.920				
DW STATISTIC	1.950				

PARAMETER ESTIMATES			
VARIABLE	ESTIMATE	T-STATISTIC	VIF
INTERCEPT	6.04	4.17	0
DUM7	-3.44	-15.37	1.6
DUM8	-2.30	-10.30	1.6
DUM9	-1.49	-6.70	1.6
DUM10	-0.65	-2.91	1.6
AIRLINE	-1.16	-10.88	2.1
PROFIT	-0.45	-2.97	4.2
UNEMP	0.79	6.39	2.8


1987 profit model
normal probability plot
ordered residual
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-0.5
-0.4
-0.3
-0.2
-0.1
0.0
0.1
0.2
0.3
0.4
0.5
expected normal value

1987 profit model

plot of residuals vs predicted



Appendix K: *Data Used in the Pay and Profit Models*

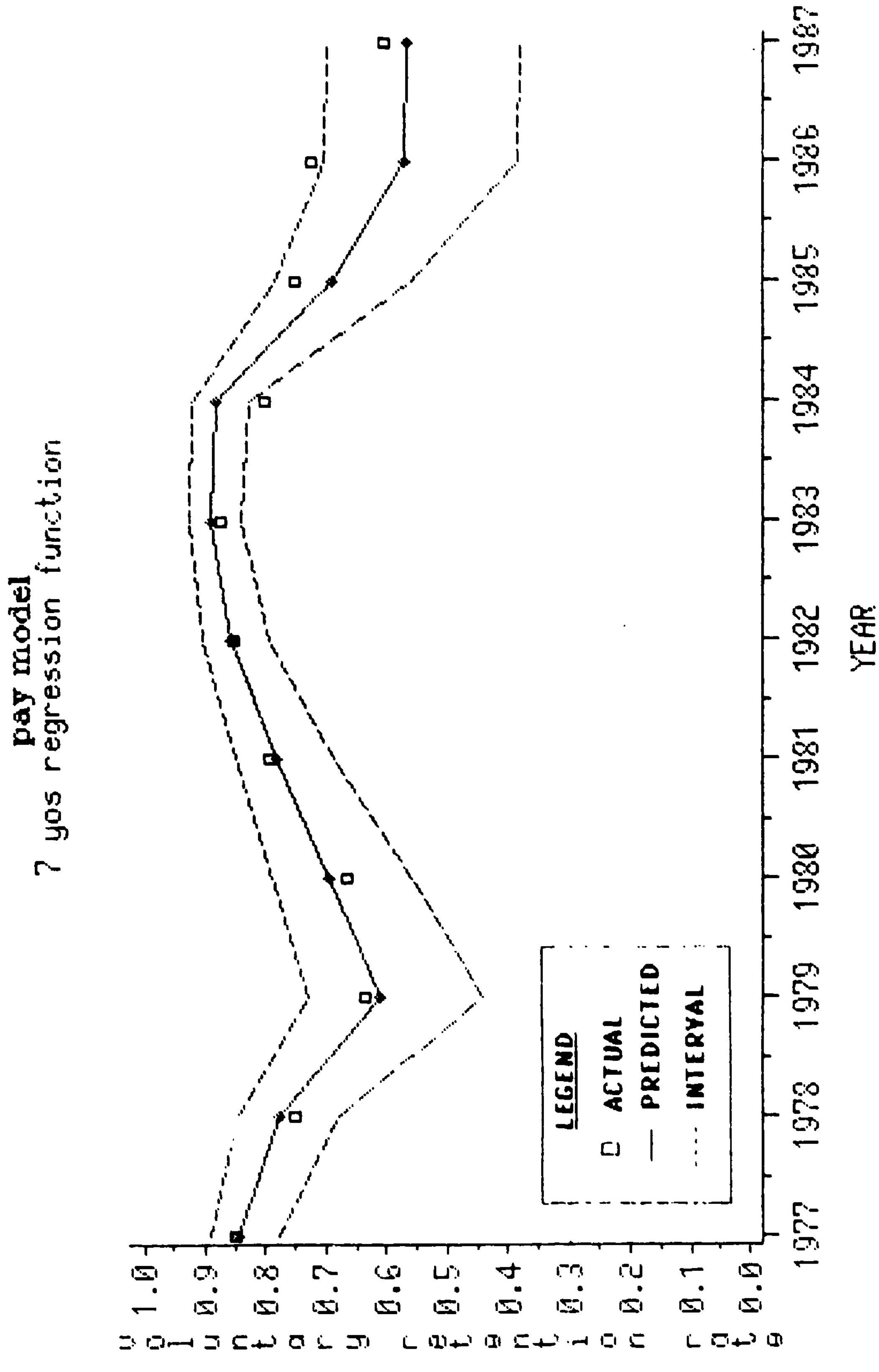
voluntary retention rates

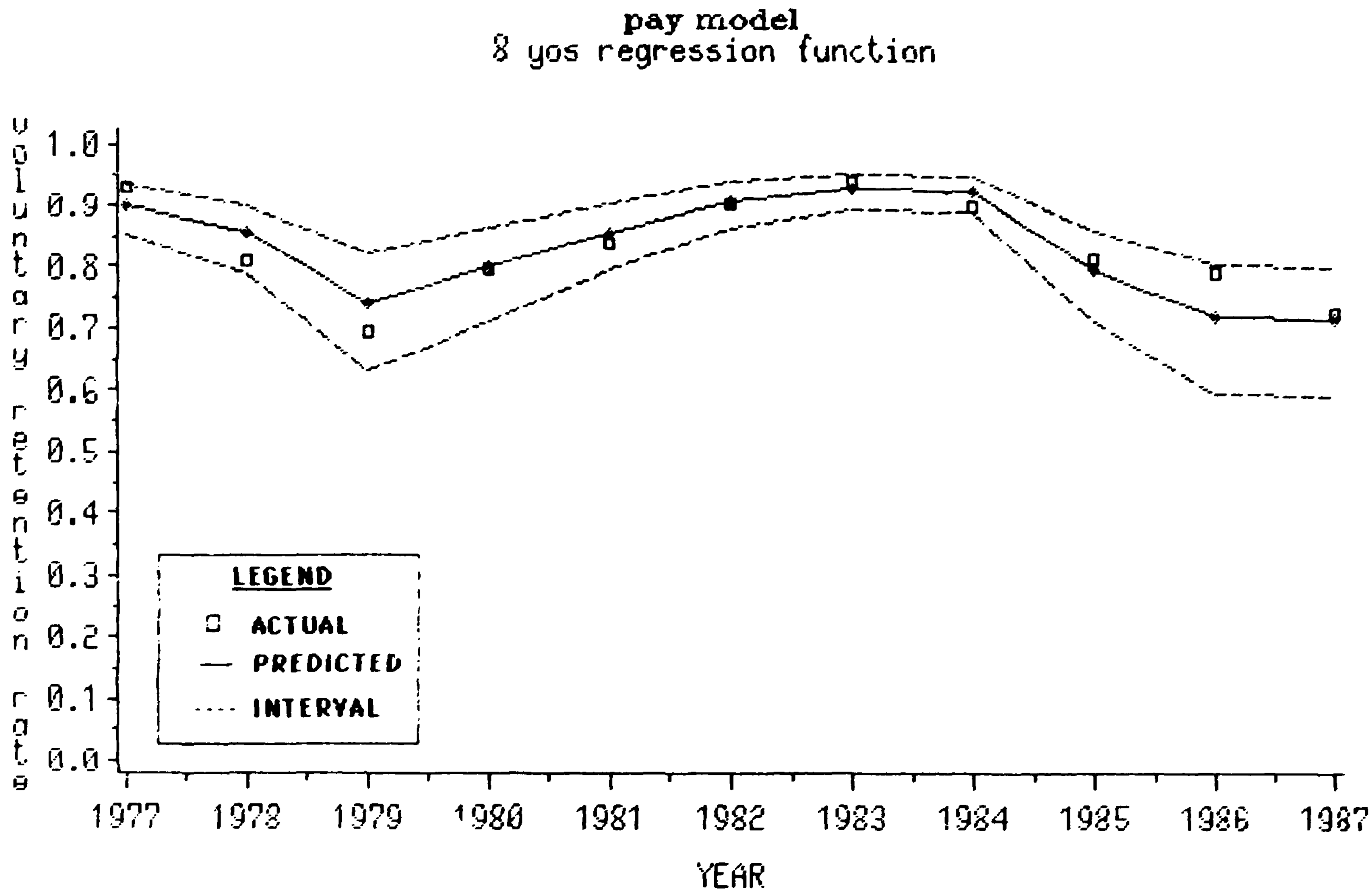
year	7yos	8yos	9yos	10yos	11yos
1977	.8480	.9313	.9513	.9515	.9684
1978	.7505	.8091	.8718	.9343	.9487
1979	.6374	.6937	.7117	.8126	.8733
1980	.6675	.7992	.8262	.8976	.9244
1981	.7925	.8418	.8900	.9156	.9585
1982	.8521	.9033	.9400	.9591	.9681
1983	.8745	.9429	.9616	.9782	.9747
1984	.8024	.9018	.9479	.9632	.9632
1985	.7494	.8143	.8634	.8782	.9334
1986	.7243	.7918	.8262	.8475	.8762
1987	.6059	.7237	.7875	.8597	.8559

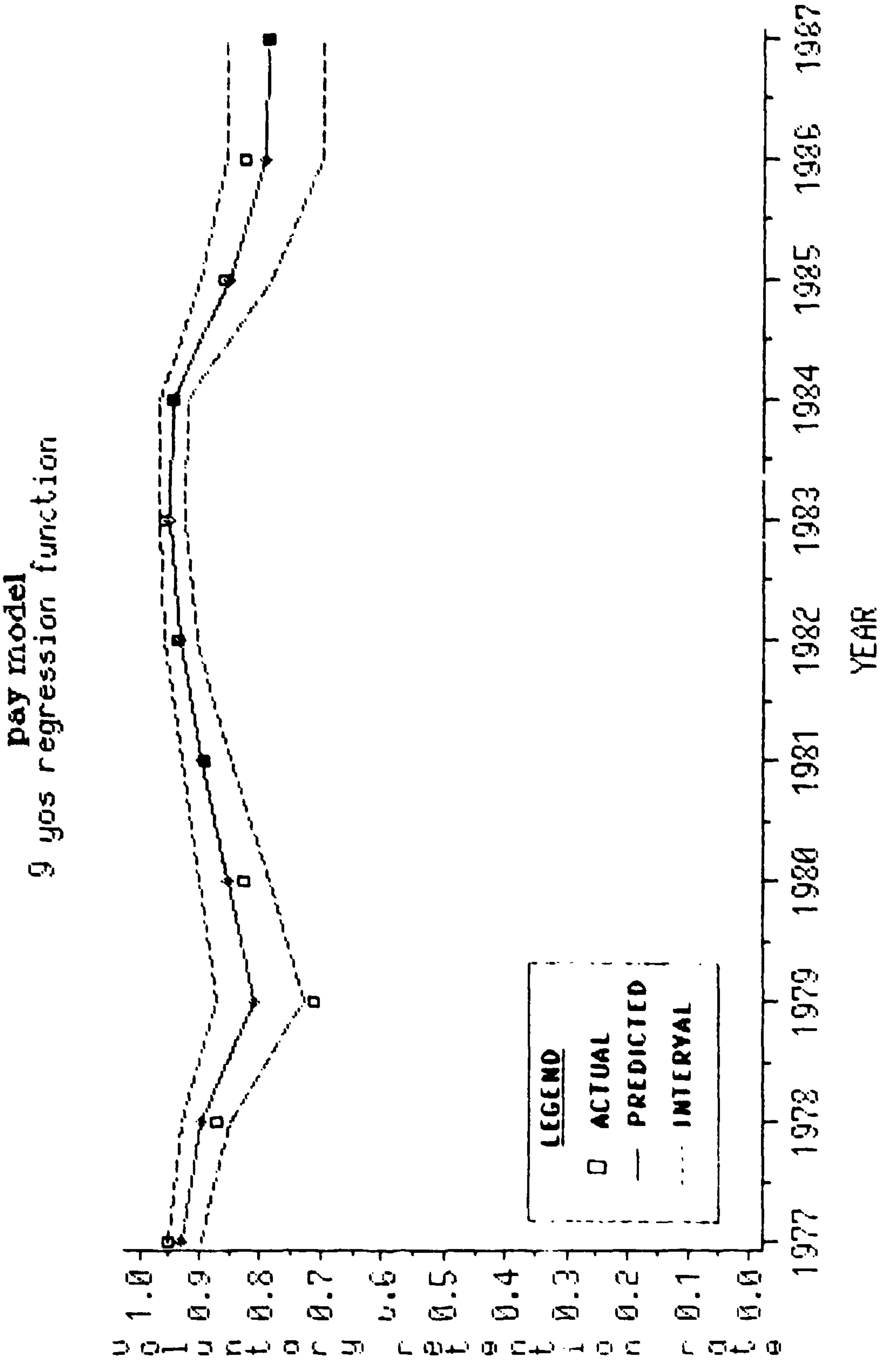
independent variables

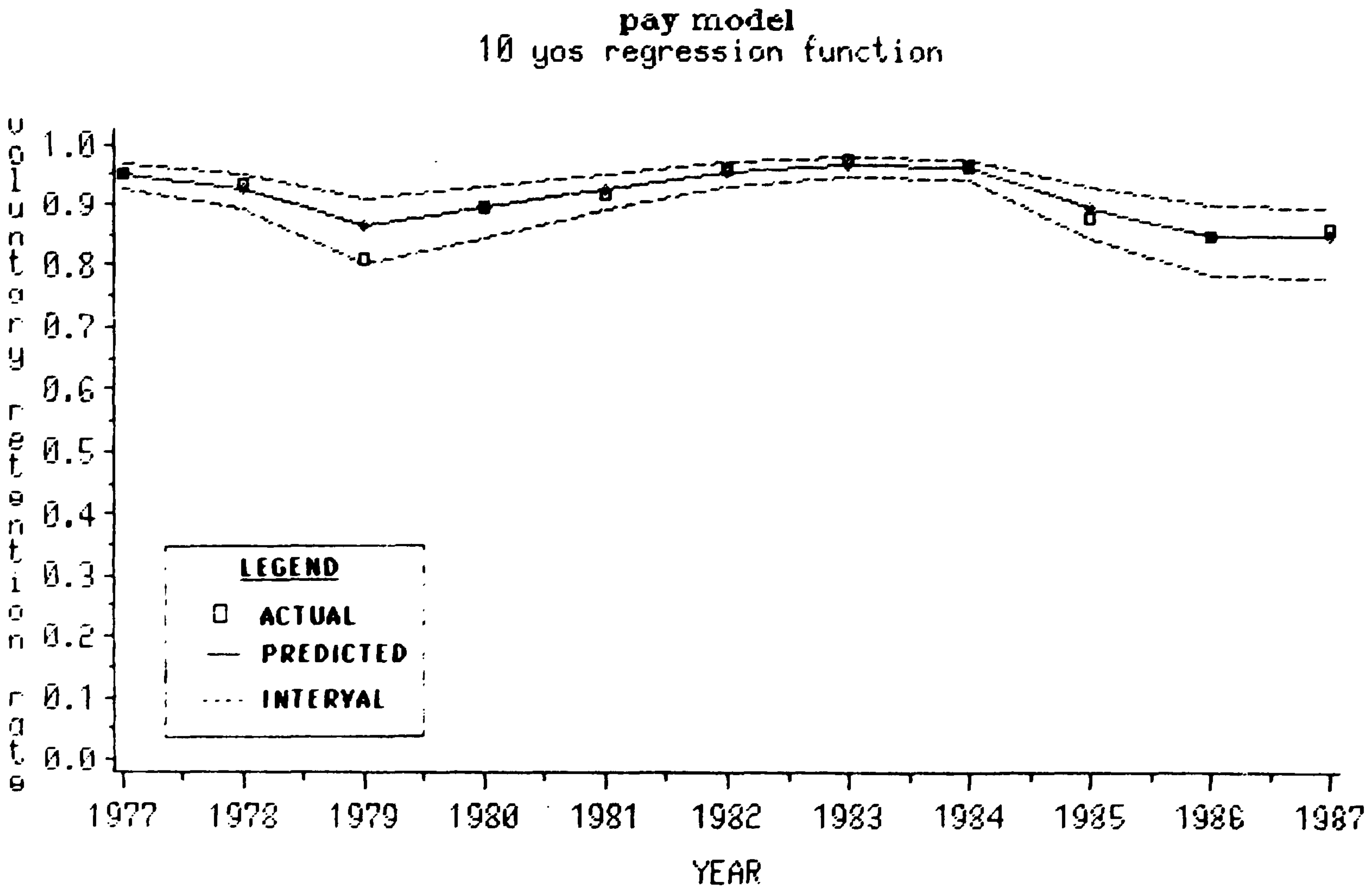
year	airline hires	pay comp	unemp rate	corporate profits
1976	449	.954	7.825	169.60
1977	1206	.957	7.325	185.75
1978	3075	.939	6.225	202.63
1979	4345	.932	5.825	219.87
1980	796	.954	6.800	188.18
1981	1319	.998	7.425	160.78
1982	881	.960	9.125	120.00
1983	1948	.946	10.125	118.88
1984	4698	.936	7.850	138.55
1985	6537	.923	7.250	123.98
1986	7334	.912	7.015	117.58
1987	6403	.902	6.425	119.76

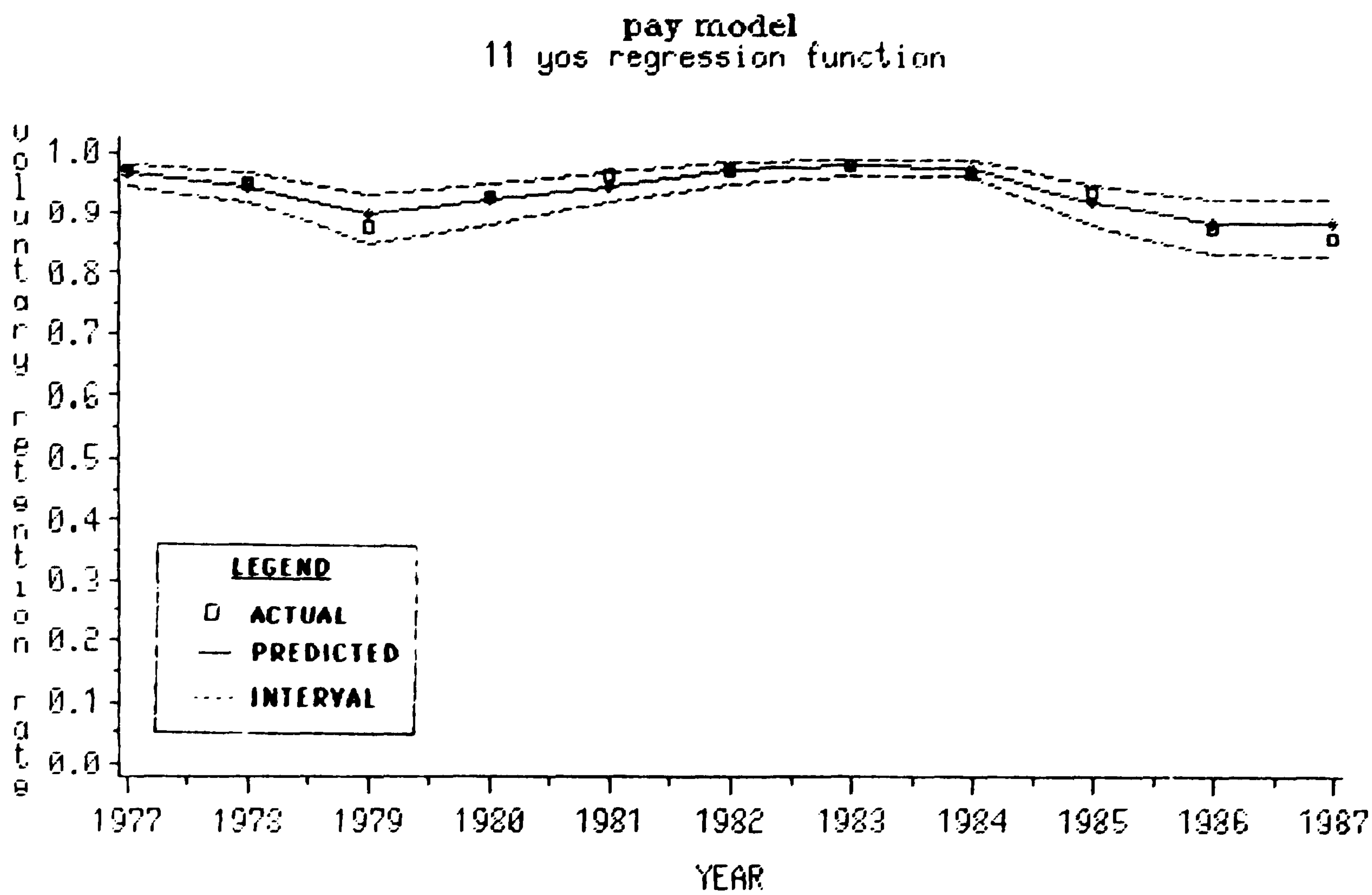
Appendix L: Plots of Actual and Predicted with Prediction Intervals- Pay Model



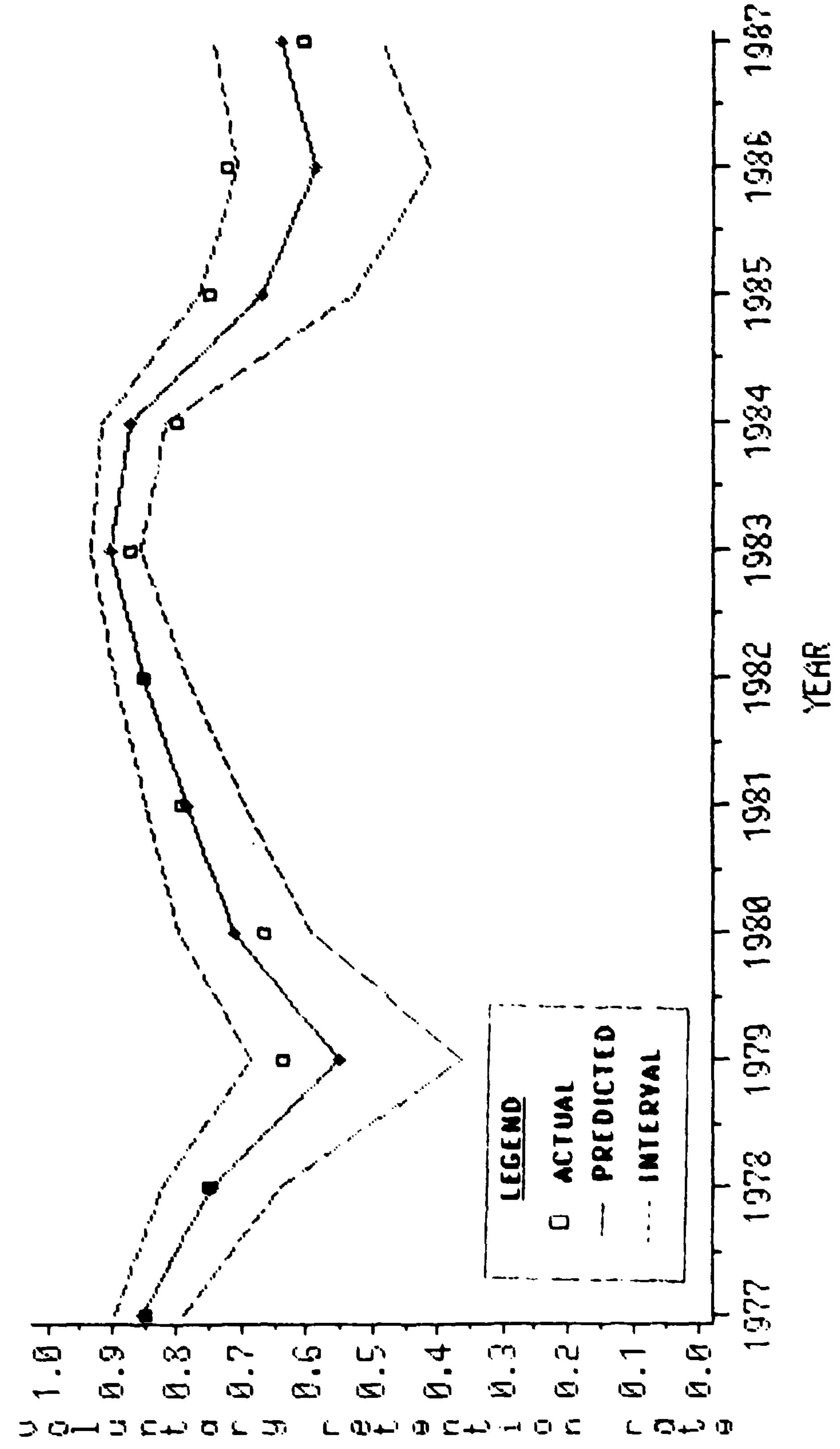

pay model
8 yos regression function
voluntary retention rate
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
YEAR
LEGEND
ACTUAL
PREDICTED
INTERVAL

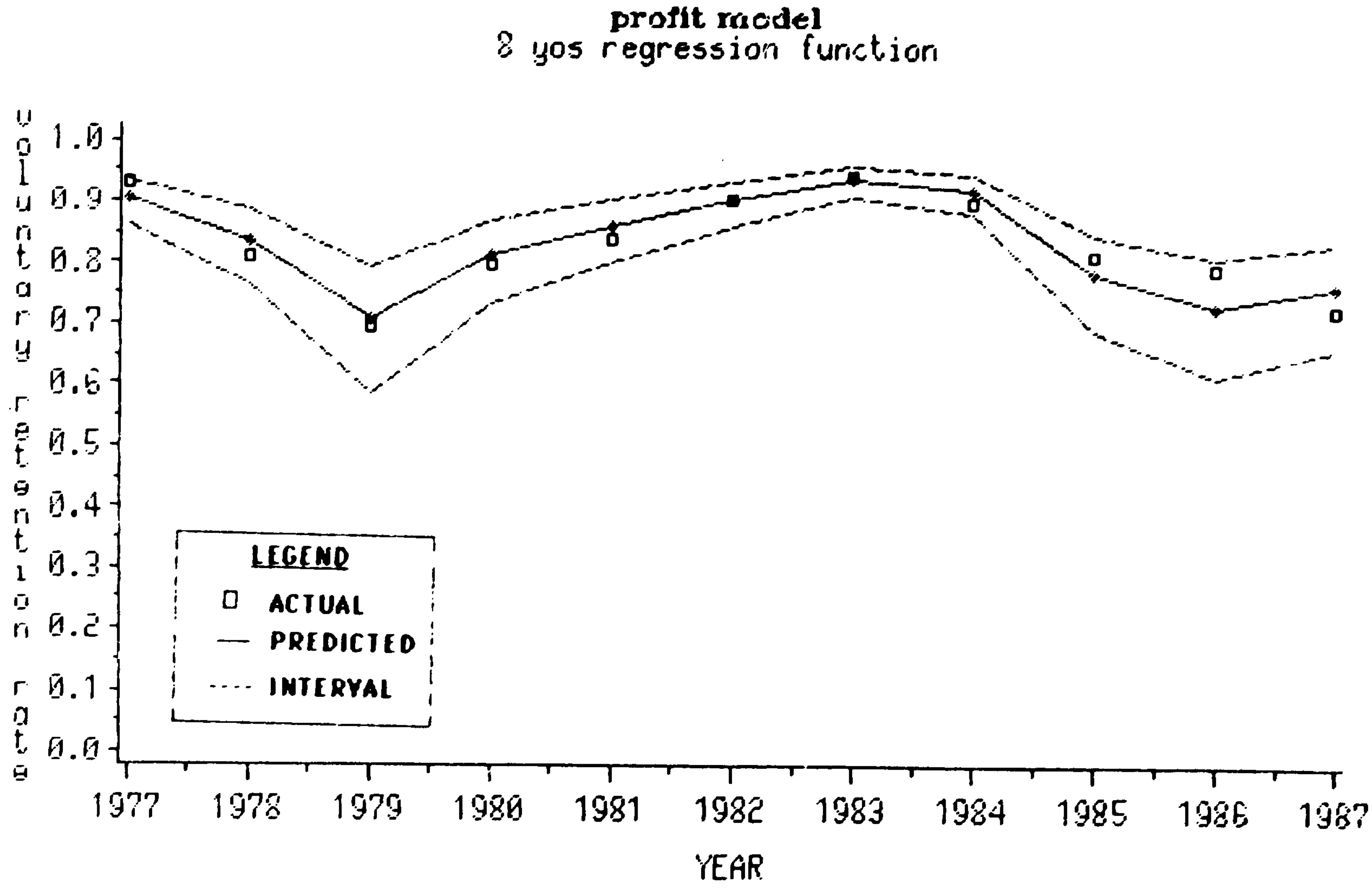

pay model
9 yos regression function
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
YEAR
LEGEND
ACTUAL
PREDICTED
INTERVAL


pay model
10 yos regression function
voluntary retention rate
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
YEAR
LEGEND
ACTUAL
PREDICTED
INTERVAL


pay model
11 yos regression function
voluntary retention rate
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
YEAR
LEGEND
ACTUAL
PREDICTED
INTERVAL

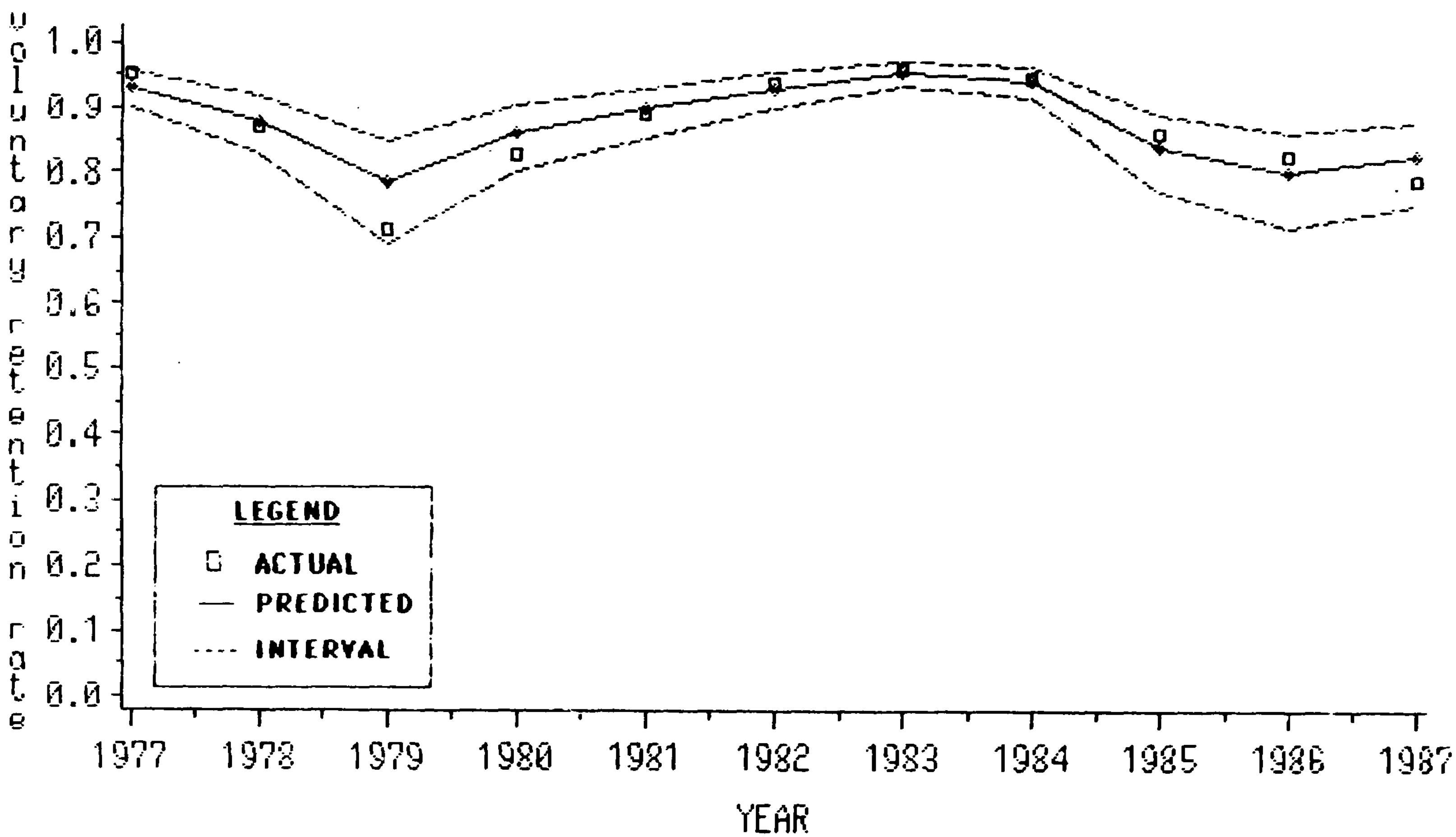
Appendix M: Plots of Actual and Predicted with Prediction Intervals- Profit Model

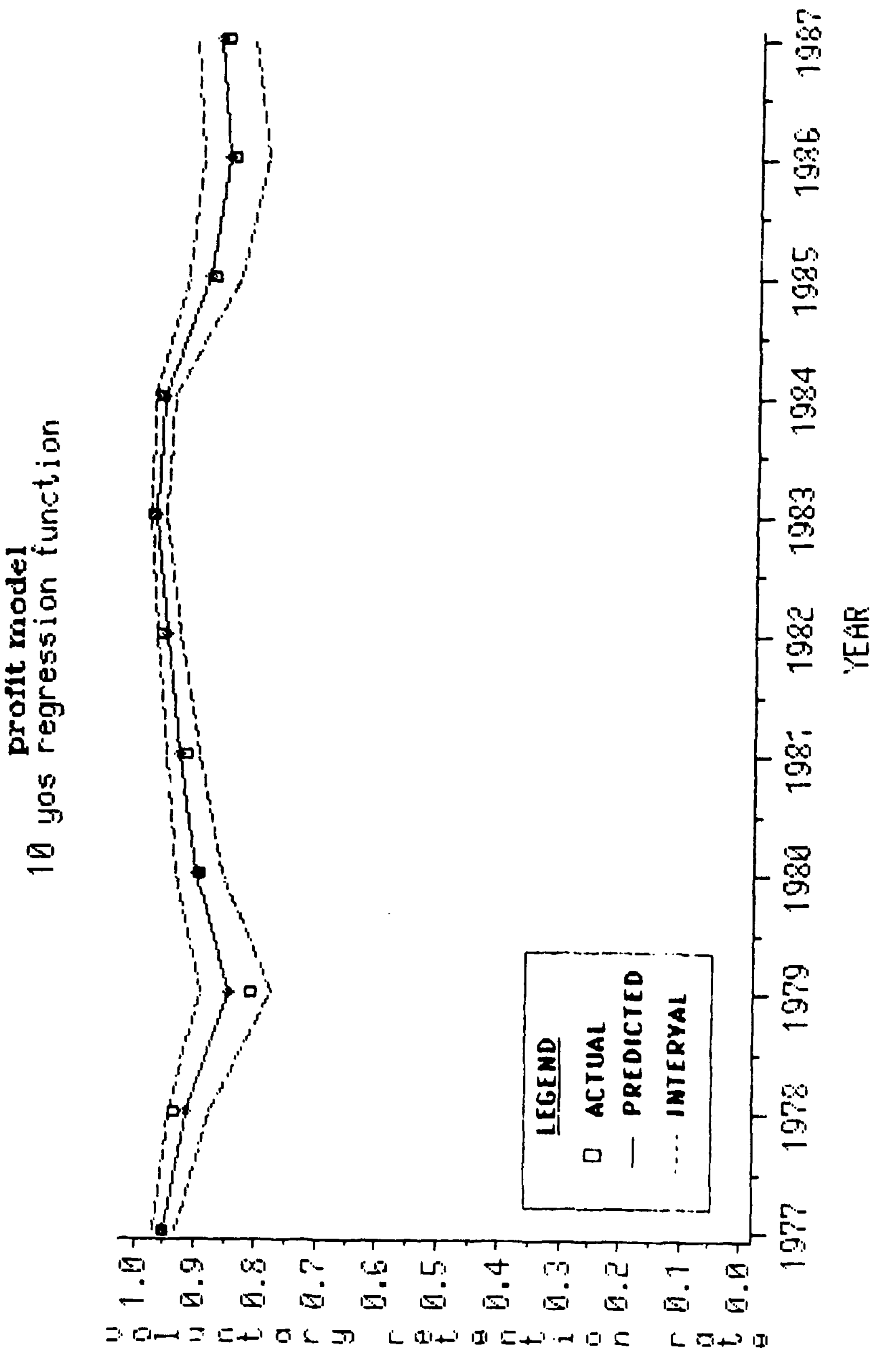




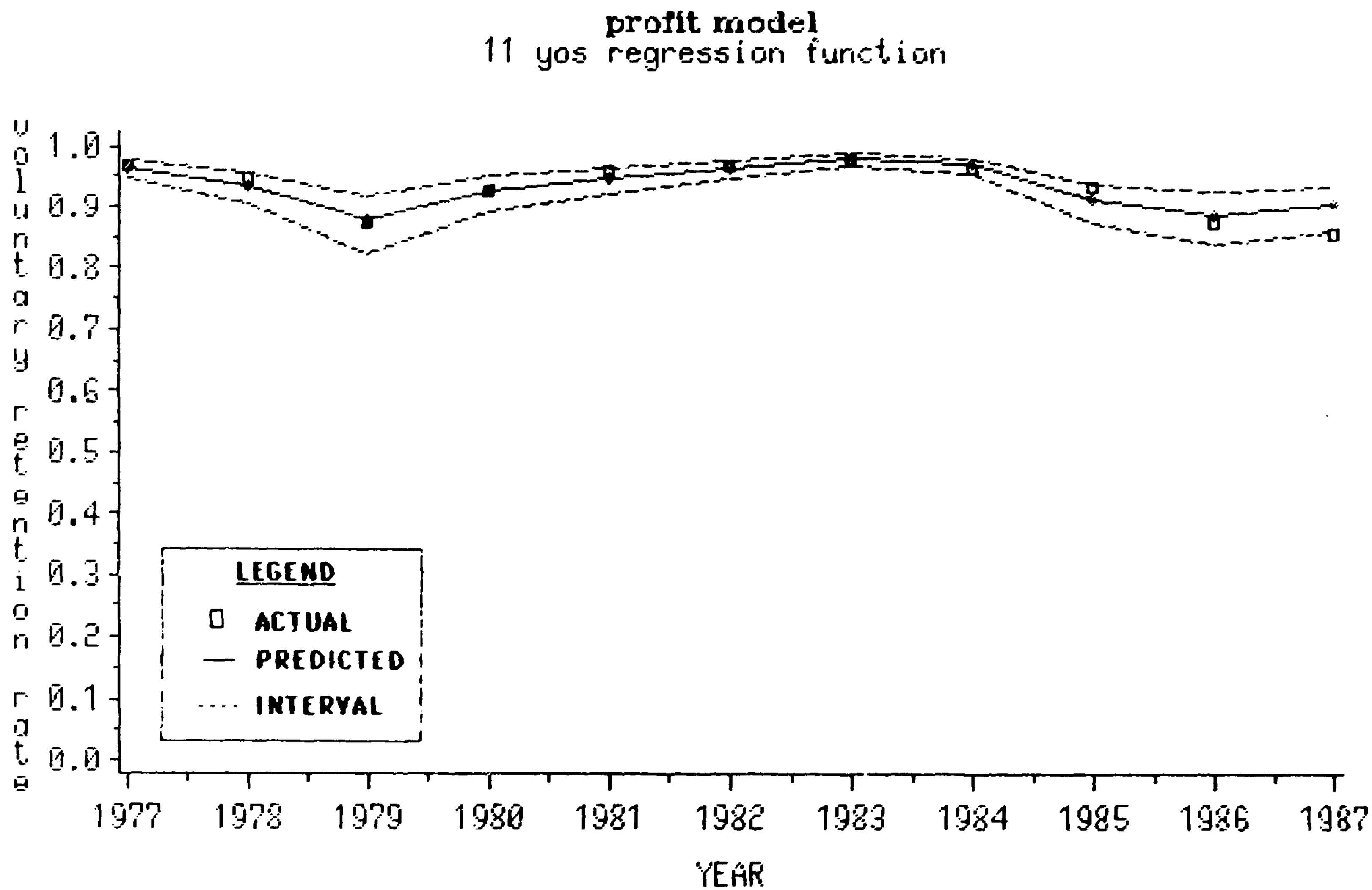
profit model
8 yos regression function
voluntary retention rate
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
YEAR
LEGEND
ACTUAL
PREDICTED
INTERVAL

profit model
9 yos regression function





profit model
10 yos regression function
voluntary retention rate
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
YEAR
LEGEND
ACTUAL
PREDICTED
INTERVAL


profit model
11 yos regression function
voluntary retention rate
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
YEAR
LEGEND
ACTUAL
PREDICTED
INTERVAL

Bibliography

1. Air Force Human Resources Laboratory, Air Force Systems Command. Contract F33615-77-C-0056 with Resources Research Corporation. Brooks AFB TX, Nov 11, 1977 (AD-A058097).

2. Air Force Human Resources Laboratory, Air Force Systems Command. Contract F33615-78-C-0049 with Resources Research Corporation. Brooks AFB TX, June 1980 (AD-A085658).

3. Air Force Human Resources Laboratory, Air Force Systems Command. Contract F33615-80-C-0021 with Resources Research Corporation. Brooks AFB TX, July 1985 (AD-A158091).

4. Bowerman, Bruce L. and Richard T. O'Connell. *Time Series Forecasting* (Second Edition). Boston: Duxbury Press, 1987.

5. Cromer, Capt DeJuan and Capt Mark R. Julicher. *An Examination of the Effects of Economic Conditions on Pilot Retention*. MS thesis , AFIT/LSSR/38-82. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, Sept 1982 (AD-A122980).

6. Gentile, Major Anthony B., *The Airlines...A Drain on Air Force Pilots (Past, Present, and Future)*. Student report No. 83-0835. Air Command and Staff College (AU), Maxwell AFB AL, June 1983 (AD-B074466).

7. Gotz, Glenn A. and John J. McCall. *A Dynamic Retention Model for Air Force Officers: Theory and Estimates: Draft Report*. Contract F49620-82-C-0018. RAND Corporation, March 1983 (AD-A149736).

8. Katnik, Lt Col, "Background Paper in Retention Measures in Perspective." Paper presented at the Retention Planning Committee meeting, Pentagon, 15 January 1986.

9. Longino, Major Dana, Analyst at the USAF Retention Division. Telephone interview. HQ AFMPC, Randolph AFB TX, 16 Oct 1987.

10. Neter, John and others. *Applied Linear Statistical Models*. Homewood IL: Richard D. Irwin, Inc., 1985.

11. Rhodes, Lt Col John D. *Pilot Retention: An Historical Analysis*. Unpublished report No. AU-AWC-86-178. Air War College, Maxwell AFB AL, 1986.

12. Sutter, Major Brian, Chief Rated Analyst for Personnel Analysis Division. Telephone interview. HQ AFMPC, Randolph AFB TX, 24 April 1987.

13. U.S. Deparment of Commerce, Bureau of Economic Analysis. *Handbook of Cyclical Indicators*. Washington: Government Printing Office, 1984.

14. U.S. Department of Labor, Bureau of Labor Statistics. *Monthly Labor Review*. Washington: Government Printing Office.

15. Vetterlein, Major Jon, Officer and Economic Analysis Branch. Telephone interview. Directorate of Personnel Plans, Pentagon, 15 July-8 Oct 1987.

16. Warner, John T. *Military Compensation and Retention: An Analysis of Alternative Models and a Simulation of a New Retention Model*. Center Research Contribution (CRC) 436. Alexandria, VA: Center for Naval Analyses, 1981 (AD-A105468).

VITA

Captain James R. Simpson was born on 2 June 1960 in Bronxville, New York. He graduated from high school in San Marino, California in 1978 and attended the United States Air Force Academy, from which he received the degree of Bachelor of Science in Operations Research in June 1982. He was commissioned a second lieutenant upon graduation. He served as a scientific analyst for the Systems Analysis and Simulation Branch at the Air Force Armament Laboratory, Eglin AFB, Florida. He served as the program manager for the Ducted Rocket Missile Simulation Program. He then entered the School of Engineering, Air Force Institute of Technology, in June 1986.

Permanent address: 1613 Chelsea Rd
Box 312
San Marino, California 91108

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) AFIT/GOR/ENS/87D-18	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION School of Engineering 6b. OFFICE SYMBOL (If applicable) AFIT/ENS	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Air Force Institute of Technology (AU) Wright-Patterson AFB, OH 45433-6583	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Personnel Analysis Div. 8b. OFFICE SYMBOL (If applicable) DPMYAF	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code) HQ AFMPC/DPMYAF Randolph AFB, TX 78150-6001	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
A METHODOLOGY FOR FORECASTING VOLUNTARY RETENTION RATES FOR AIR FORCE PILOTS

12. PERSONAL AUTHOR(S)
James R. Simpson, Capt, USAF

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
MS Thesis	FROM ______ TO ______	1987 December	102

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP
05	09	
12	03	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Regression analysis, Retention, Officer Personnel, Military Separation

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Thesis Advisors: Joseph R. Litko, Major, USAF
Assistant Professor of Operations Research

Bruce Christensen, Lt Col, USAF
Deputy Head Dept of Logistics Management

Approved for public release: IAW AFR 190-17.
LYNN E. WOLAVER
Dean for Research and Professional Development
Air Force Institute of Technology (AFC)
Wright-Patterson AFB OH 45433

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Joseph R. Litko, Ph.D.	22b. TELEPHONE (Include Area Code) (513) 255-3362 22c. OFFICE SYMBOL AFIT/ENS

19. The purpose of this study is to develop a model that more acurately forecasts voluntary retention rates in the short term for Air Force pilots. Specifically, the model consists of appropriate and available predictors used to compute one year ahead forecasts of voluntary retention rates for Air Force pilots with seven through eleven years of service. The types of predictors collected for study were indicators of the strength of the economy, indicators of the growth of the airline industry, and indicators of the relative wage difference between the military and the civilian labor force. Classical regression analysis was used to predict the pilot retention rates on the basis of the predictor variables studied.

A logarithmic transform of the dependent variable was used to stabilize the variance of the error terms. The criteria established for selecting the best model were model performance, prediction potential, and explanatory significance. The best model included the following independent variables: indicator variables for the year of service groups, a variable for the annual number of new airline pilot hires, the unemployment rate lagged one year, and a pay compensation measure lagged one year. Thus, estimates were required only for the airline hires predictor in order to forecast pilot retention rates.

Validation tests were performed on the best model for years 1986 and 1987. In each test, the 90 percent prediction intervals covered the actual pilot retention rate for each year of service group. Among the recommendations provided to improve the accuracy of the pilot retention rate forecasts was to improve the accuracy of the airline hire forecasts and to find other significant, leading indicators of pilot retention. (Theses)